Christoffer Kauppinen

Fabrication and characterization of plasmonic nanogratings on indium gallium nitride quantum wells

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 10.3.2014

Thesis supervisor:

Prof. Markku Sopanen

Thesis advisor:

D.Sc. (Tech.) Sami Kujala

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

Author: Christoffer Kauppinen

Title: Fabrication and characterization of plasmonic nanogratings on indium gallium nitride quantum wells

Date: 10.3.2014 Language: English Number of pages:6+58

Department of Micro- and Nanosciences

Professorship: Optoelectronics Code: S-104

Supervisor: Prof. Markku Sopanen

Advisor: D.Sc. (Tech.) Sami Kujala

In this work a process for fabricating subwavelength silver gratings on GaN was developed, using azo-polymer surface relief gratings as etch masks. The azo-polymer gratings were fabricated using laser interference lithography. The period of the gratings was 255 nm. The patterned areas were approximately 4 cm^2. The purpose of the gratings was to couple energy from an InGaN quantum well inside the GaN to surface plasmons, and to couple these surface plasmons to radiative modes.

Two-dimensional arrays of gold spheres were fabricated with the same technique. The diameter of the spheres was approximately 40 nm. These were used to grow GaAs nanopillars on GaAs using metalorganic vapour phase epitaxy.

Surface plasmon coupling experiments were also performed on the processed GaN samples. An optical setup was built for this purpose. The results from the coupling experiments were inconclusive, although the experimental data agrees qualitatively with theory.

Keywords: Surface plasmons, InGaN, LED, azo-polymer, laser interference lithography, nanostructure

AALTO-YLIOPISTO
SÄHKOTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Christoffer Kauppinen

Työn nimi: Plasmonisten nanohilojen valmistus ja karakterisointi indiumgalliumnitridi kvanttikaivojen päälle

Päivämäärä: 10.3.2014 Kieli: Englanti Sivumäärä:6+58

Mikro- ja nanotekniikan laitos

Professuuri: Optoelektroniikka Koodi: S-104

Valvoja: Prof. Markku Sopanen

Ohjaaja: TkT Sami Kujala

Tässä työssä kehitettiin prosessi hopeahilojen valmistamiseksi galliumnitridin päälle, käyttäen atsopolymeerihiloja etsausmaskeina. Atsopolymeerihilat valmistettiin laserinterferenssilitografialla. Hopeahilojen periodi oli 255 nm. Kuvioidut näytteet olivat pinta-alaltaan noin 4 cm^2. Hopeahilojen tarkoitus oli kytkeä energiaa galliumnitridin sisällä olevasta indiumgalliumnitridikvanttikaivosta pintaplasmoneihin, ja kytkeä energiaa pintaplasmoneista säteileviin moodeihin.

Kaksiulotteisia kultapallomatriiseja valmistettiin samalla tekniikalla. Kultapallojen halkaisija oli noin 40 nm. Näitä käytettiin galliumarsenidinanopilareiden kasvattamiseksi galliumarsenidin päälle käyttäen metallo-orgaanista kaasufaasiepitaksiaa.

Pintaplasmonien kytkeytymiskokeita suoritettiin prosessoiduille galliumnitridinäytteille. Optinen mittausjärjestely rakennettiin tätä varten. Plasmonista kytkeytymistä ei varmuudella voitu osoittaa, vaikka kytkeytymiskokeiden tulokset sopivat kvalitatiivisesti teorian kanssa yhteen.

Avainsanat: Pintaplasmoni, InGaN, LED, atsopolymeeri, laserinterferenssilitografia, nanorakenne

Preface

I would like to thank my teachers for the knowledge, that enabled this work to happen. Especially I would like to mention Simo Korpela, Jarmo Perttula, Petri Salo, Taina Kurki-Suonio, Antti Hakola, Matti Stenroos, Keijo Nikoskinen and Matti Taskinen.

My father Arto Kauppinen deserves thanks for igniting my interest in matters scientifical. His ability to explain complex topics to a child is beyond wondrous.

I would like to thank Ville Pale for his idea of using azo-polymers in this work. The knowledge of Aleksandr Kravchenko proved to be invaluable, therefore I would also like to thank him for his advice.

For patience, support and good guidance I would like to thank my supervisor Prof. Markku Sopanen and my instructor D.Sc. Sami Kujala.

Otaniemi, 9.3.2014

Christoffer J. Kauppinen

Contents

Symbols and abbreviations

Symbols

c_0	the speed of light in vacuum
$\mathbf{E}$	electric field
f	frequency
$\mathbf{H}$	magnetic field density
$\hbar$	Planck's constant divided by 2π
i	the imaginary unit
$\mathbf{k}$	wave vector
k_B	Boltzmann constant
m^*	effective mass
n	refractive index
$\mathbf{p}$	electric dipole moment
q	electric charge
T	temperature
V	electric potential

β	propagation constant
ϵ	permittivity
θ	incidence angle
λ	wavelength
ρ	plasmon density of states
ψ	wave function
ω	angular frequency
ω_p	plasma frequency

Abbreviations

AFM	atomic force microscopy
EBL	electron beam lithography
IMI	insulator-metal-insulator
MOVPE	metalorganic vapour phase epitaxy
NA	numerical aperture
PL	photoluminescence
PR	polarization ratio
QW	quantum well
RIE	reactive ion etching
SEM	scanning electron microscope
SP	surface plasmon
SQW	single quantum well
TM	transverse magnetic
UV	ultraviolet
XRR	X-ray reflection

1 Introduction

The most important use for electric power has been, and still is, to enlighten the world, literally and also figuratively. The electric light has enabled easy and safe lighting for countless of people. It has made possible to read far into the night, and thus helped to civilize the world. Today, a fifth of electricity is consumed by lighting [1]. Even small efficiency increases in lighting devices can deliver massive savings in electricity consumption.

Traditional incandescent and fluorescent lights have low luminous efficacies (lumens produced per watt), typically $1\,\mathrm{lm\,W^{-1}}$ and $20\,\mathrm{lm\,W^{-1}}$, respectively [2]. Efficiency and luminous efficacy are linked so they are often used synonymously. White light-emitting diodes (LEDs) have a typical efficiency of $80\,\mathrm{lm\,W^{-1}}$ [3]. Therefore, LEDs have the potential for large savings in electricity.

In white LEDs made with InGaN quantum wells (QWs), blue light is emitted and the blue light excites different phosphors on the LED. The phosphors then emit yellow light [2]. Thus white light can be generated with LEDs. InGaN LEDs are mostly GaN with only the QW containing indium. White light can also be generated with LEDs by combining red, green and blue LEDs in a single light source. Although commercial blue LEDs have a higher luminous efficacy than incandecent light bulbs, some fluorescent tubes still have better efficiency [3]. However, LEDs have many superior qualities compared to fluorescent lighting, notably their lifetime is far longer [2]. High performance blue LEDs would therefore be very useful.

The InGaN material itself is the main reason why white LEDs have low efficiency. The high refractive index ($n = 2.41$) of gallium nitride (GaN) prevents most of the light from radiating away from the air/GaN interface due to total internal reflection. This leads to a low external quantum efficiency. External quantum efficiency describes here how efficiently the used energy is converted to light. The losses originating from the high probability of total internal reflection have been reduced by multiple techniques.

Index matching has been the traditional way to increase external quantum efficiency in InGaN LEDs. The semiconductor crystal is buried in a plastic, whose refractive index is between air and GaN. This enables more light to radiate, due to the refractive index contrast being smaller. The index matching plastic is usually the plastic cap in packaged LEDs. The cap can also be dome-shaped to give all rays emanating from the semiconductor crystal a near normal air/plastic interface. Index matching can be, and is, used in conjunction with more modern techniques.

Microroughening is another method to overcome some of the losses caused by total internal reflection. Studies have found that a microroughened GaN surface on an InGaN LED can increase the external quantum efficiency [4, 5]. Power conversion efficiency increases over 60% have been reported [4]. The GaN/air (or GaN/plastic) interface is etched to achieve uniform nanoscale roughness on the interface. Scattering from the roughness

enables the emitted photon to experience multiple chances of transmission. This permits more light to escape the crystal.

Photonic crystals have also been investigated as a possible way to increase the external quantum efficiency of InGaN LEDs [6, 7]. Photonic crystals are media, where the permittivity changes periodically [8]. This allows only certain wavelengths to propagate in certain directions of the structure [8]. They are often made of two materials, for example silicon dioxide and GaN [6]. The photonic crystal in or on (see Ref. 7) the LED can couple light from guided modes to air [6].

Higher external quantum efficiencies are also being pursued by using surface plasmons (SPs) [9–12]. The idea is to couple the energy of the photon to a SP [13]. Then certain structures are used to cause the SP to radiate. This way the total internal reflection problem can be bypassed. The most common structures in InGaN related SP coupling are different kinds of metallic gratings [9–11]. Succesful InGaN SP LEDs have been demostrated, and the research is ongoing [9, 10, 12].

The main interest in this thesis is to study the feasibility of using azo-polymer masks to fabricate nanoscale plamonic gratings to couple the recombining electron hole pairs of the InGaN QW directly to surface plasmons, see Chapters 4 and 6 or Ref. [13]. These plasmons are then coupled to radiating light. It may be possible to increase the external quantum efficiency this way. The same principle for improvement can be used for other types of QW LEDs. The method of manufacturing the gratings was inspired by studies conducted by A. Kravchenko [14, 15]. The method used here is somewhat similar, see Chapter 5. Input to the manufacturing process was received from A. Kravchenko as well. The silver gratings chosen to be manufactured are 150 nm wide square profiled gratings with a period of 250 nm. Similar gratings by E. Homeyer *et al.* were fabricated by electron beam lithography (EBL) [11].

This unorthodox patterning technique is used for several reasons. First, producing 150 nm wide silver gratings with 100 nm spacing is extremely challenging with conventional UV (ultraviolet) lithography as this requires subwavelength accuracy. The second reason is that although EBL can produce comparably small details, the fabrication of large area gratings with EBL is very slow [14]. The method employed in this work can produce very large 2 cm times 2 cm gratings. Finally, the electron beam causes serious damage to the InGaN QWs [16]. The damage to the QWs is limited in depth [16]. The thicker the so called capping layer of GaN above the InGaN QW is, the less the QW is damaged. However, the possible plasmonic coupling drops exponentially as a function of QW depth [13], thereby making this method of protection problematic.

Several different charcterization methods are used in this work. Atomic force microscopy (AFM) and scanning electron microscopy (SEM) were used to study the surface structure of the samples. SEM was also used to study the material distribution on the samples, because metals and dielectrics can be easily distinguished in SEM. X-ray reflection (XRR) was used to study the layer thicknesses in the epitaxially grown GaN and InGaN layers.

Photoluminescence (PL) was used to study the band gap energy of the QWs in the samples as well as the plasmonic coupling.

This thesis first reviews the theory of semiconductors and how InGaN QW LEDs operate in the second chapter. The third chapter focuses on the GaN/InGaN material system and the fourth explains what SPs are and how they can be harnessed in incresing the external quantum efficiency of QWs. The fifth chapter deals with the manufacturing of the gratings. The sixth chapter presents some fabricated structures and the light extraction measurement. Finally the seventh chapter concludes and summarizes this work.

2 Theoretical background of quantum well LEDs

In this chapter the general characteristics of semiconductors are discussed. Several interesting topics fundamental to the understanding of semiconductors are examined. The bandgap, holes as carries of current, the p-n junction, the quantum well and the light-emitting diode (LED) are analyzed. The purpose of this section is, through these topics, to explain how the QW-LED operates.

2.1 Properties of semiconductors

Semiconductors are typically crystalline elements that are characterized by their conductivity, which falls between conductors (metals) and insulators (the rest). The poor conductivity of insulators stems from the fact that in order for current to flow, a partially filled electron band is needed [17]. The difference between insulators and semiconductors is that insulators have a larger bandgap than semiconductors (the classification is indeed somewhat artificial but useful). The band gap is the difference in energy between electron bands (explained later), and it is a dominating characteristic of every semiconductor material or device. Many high band gap semiconductors, like GaN the semiconductor used in this work, can be classified as insulators.

The conductivity of pure semiconductors can be modified by several orders of magnitude. This happens by introducing impurity elements, called dopants, into the crystal. These provide donor or acceptor atoms whose electrons or holes (explained later) can be thermally excited to provide the crystal with more charge carriers to conduct current. Doped semiconductors are called extrinsic semiconductors. When a semiconductor is not doped, and otherwise is free from impurities, it is called an intrinsic semiconductor. In intrinsic semiconductors the conduction is due to thermally (or otherwise) excited electrons in the conduction band, and the holes that they leave behind in the valence band. In equilibrium the concentration of electrons in the conduction band n and the concentration of holes in the valence band p are equal. At room temperature the intrinsic concentration is very small and not very significant.

2.2 Electrons in a crystal – the origin of the band gap

The electrons in a crystal are described by the time-independent Schrödinger equation

$$\left(\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r})\,. \tag{1}$$

The time independence assumption is called the adiabatic approximation [18]. The approximation can be done because the movement of the nuclei is very slow compared to the electrons, due to the minuscule mass of the electrons [17]. The electrons assume a new position almost instantly when a nucleus moves and to a good degree the electron system remains at the ground state. Therefore, the time independence is a valid assumption [17].

Now when time independence of the Schrödinger equation for electrons in the crystal is established, other approximations are needed for understanding the band gap. Perfect periodicity of the crystal lattice is one of them [17]. Perfect periodicity means also that the crystal is infinite in all dimensions so that no edges disturb the periodicity. More information about crystal lattices can be found in Ref. 17. The single electron approximations is also needed. Electron-electron interactions are neglected, and a single electron thus sees only a periodic potential

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{r_n}) \tag{2}$$

of the lattice atoms [17]. The vector describing the lattice periodicity is

$$\mathbf{r_n} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3,\ n_i \in \mathbb{Z}, \tag{3}$$

where $\mathbf{a}_i$ are the lattice vectors. The solution for the Schrödinger equation with this kind of potential can be written as a modulated plane wave

$$\psi_\mathbf{k} = u_\mathbf{k}(\mathbf{r})\, e^{i\mathbf{k}\cdot\mathbf{r}} \quad . \tag{4}$$

Here $\mathbf{k}$ is the reciprocal lattice vector, also called the crystal momentum. The reciprocal lattice emerges naturally when considering diffraction from crystals, but omitting a long discussion about diffraction in crystals [17] the reciprocal lattice is determined by the following expression

$$\mathbf{G}\cdot\mathbf{r_n} = 2\pi m,\ m \in \mathbb{Z}, \tag{5}$$

where $\mathbf{r_n}$ is a lattice vector (a lattice site in other words). The modulating function $u_\mathbf{k}(\mathbf{r})$ has the periodicity of the lattice, namely $u_\mathbf{k}(\mathbf{r}) = u_\mathbf{k}(\mathbf{r} + \mathbf{r_n})$.

This solution is called Bloch's theorem and wavefunctions having such properties are called Bloch waves, see Fig. 1. For Bloch waves it can be shown [17] that

$$\psi_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) = \psi_\mathbf{k}(\mathbf{r}) \tag{6}$$

where $\mathbf{G}$ is an arbitrary reciprocal lattice vector. Thus the energy eigenvalues of the Schrödinger equation have the same periodicity. Therefore the wavefunction needs to be known only in the first unit cell in k-space, the so called Brillouin zone. The Brilloun zone wavefunction is then periodically repeated in the crystal.

2.2.1 Energy bands

Free atoms have electrons in their respective orbits with discrete energies. When atoms form a crystal, however, the Hamiltonian of the single atom is altered by the other particles in the crystal (mainly by the nearest neighbour atoms). This causes the formerly discrete energy levels to broaden into quasi-continous "bands", where the electron energy is determined by its place in the reciprocal space, *i.e.*, k-space.

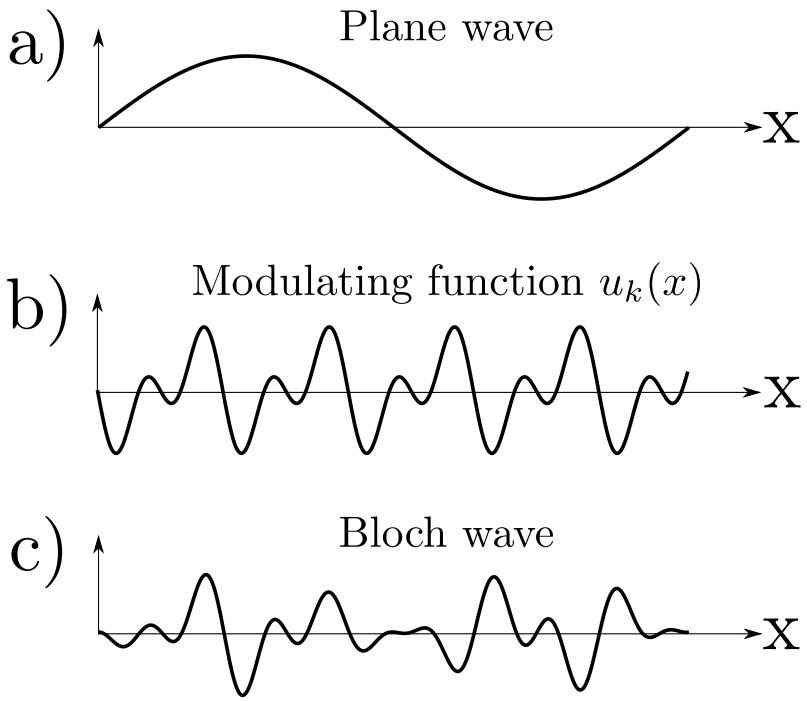

Figure 1: Simple visualization of Bloch waves in a crystal. a) A plane wave, specifically its real part, depicted propagating along an arbitrary direction marked x, b) the modulating function which has the periodicity of the lattice and c) the resultant Bloch wave.

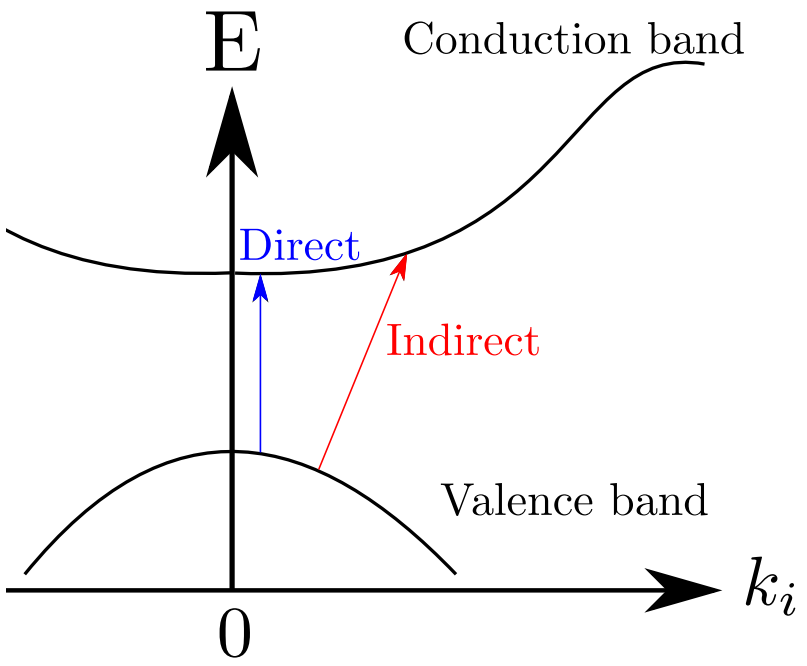

Figure 2: Illustration of the band structure of a direct band gap semiconductor. The blue arrow represents a direct transition of an electron. The red arrow represents an indirect transition.

Two simple models predict the formation of energy bands. Both are approximations of the actual situation. Approximations are needed because a cubic centimeter of solid material contains about 10^{23} atoms and solving the Schrödinger equation for this many atoms is impossible. Both models are often called by their respective approximation.

In the nearly-free electron approximation it is assumed that the electrons experience a neglibly small periodic potential. At the edges of the crystal the potential is assumed infinite. Furthermore, it is assumed, that the electrons do not interact (a one-electron approximation). From these starting points the allowed bands and the forbidden bands emerge. Forbidden means that with a certain k-value there exists an energy gap where no electron states may appear.

The tight-binding approximation is almost the opposite to the nearly free-electron approximation. It starts from the fact that electrons in the lowermost orbitals are strongly localized in space. The energy levels are calculated for only one electron with a Hamiltonian that is the free-atom Hamiltonian plus the influence of the nearby atoms. The tight-binding approximation also predicts the existence of energy bands. A discussion about the tight binding approximation is given in Ref. 17.

It is remarkable, that two totally different theoretical models predict the same band structure. This is a strong hint that band structure indeed exists in semiconductors. Although the band structure can be predicted from these models, the width of the band gap cannot be calculated.

Transitions between the bands are extremely important in optoelectronics. The smallest energy difference between the valence band and the conduction band (in k-space) is called the gap energy E_{gap}. If the smallest gap is between two points having the same k-value the semiconductor is called a direct band gap material, if not it is called an indirect band gap material. Silicon is an example of an indirect band gap material and gallium nitride is an example of a direct band gap material. A direct band gap Ek-plot is shown in Fig. 2.

Two types of band to band transitions exist, direct transitions and indirect transitions. The most important direct transition concerning this work is the direct optical transition. An example of an optical transition is when a conduction electron emits a photon when recombining with a valence band hole. Holes are described in Chapter 2.2.2 . Indirect transitions require that an electron gains or loses momentum. This can happen when the electron interacts with a phonon (a quantum of heat). This is a three-particle process, and therefore, much more unlikely to happen than the direct transition. Therefore indirect band gap materials are not so efficient emitters. In Fig. 2 the red transition would require the electron to gain momentum from a phonon. Compared to phonons, photons carry several orders of magnitude smaller amounts of momentum and the minuscule amount of momentum transferred by photons is usually ignored in optical transitions [17].

In addition to optical (radiative) recombination, there are two major recombination processes in crystals: defect and surface recombination [19]. Recombination via defects

happens from impurity states in the forbidden zone. Recombination can be radiative, that is excess energy is radiated as a photon, or non-radiative where the excess energy is probably lost by emitting phonons [17]. Surface recombination is due to the huge amount of impurity states and dangling bonds near the surface [17].

2.2.2 Carrier effective masses and holes

To continue being able to describe electrons as semiclassical particles, their mass needs to be replaced with an effective mass tensor m^*. Effective mass contains the contributions of the crystal to the response of the electron when an outside force tries to move it. Typically effective masses in crystals are lower than in vacuum. This means that it is easier to move an electron in a crystal, than in vacuum! If the effective masses in all the principal directions are the same the effective mass reduces to

$$m^* = \hbar^2 \left(\frac{d^2 E}{dk^2} \right)^{-1}, \tag{7}$$

where $\hbar = \frac{h}{2\pi}$ and h is the Planck constant, k is the crystal momentum and E is the energy [17].

The energy of an electron in the first Brillouin zone is

$$E(\mathbf{k}) = \frac{\hbar^2}{2} \left(\frac{{k_x}^2}{m_x} + \frac{{k_y}^2}{m_y} + \frac{{k_z}^2}{m_z} \right), \tag{8}$$

where the m_i corresponds to effective masses in directions i [19]. In zinc-blende semiconductors, such as GaAs or zinc-blende GaN, this can be further simplified because the constant energy surfaces near the X and L points are ellipsoids, and the energy simplifies to

$$E(\mathbf{k}) = \frac{\hbar^2}{2} \left(\frac{{k_l}^2}{m_l} + \frac{2{k_t}^2}{m_t} \right), \tag{9}$$

where the subscript l denotes longitudinal and t transverse direction of the ellipsoid in question [19]. This means the effective mass can exhibit anisotropy.

The effective mass can also have a negative value (or even an indeterminate value!) depending on the concavity of the energy-momentum curve. A carrier that has a negative mass and negative charge is interpreted as a carrier with a positive mass and positive charge [17, 19]. This is because under the effect of an electric field these particles behave exactly the same. These positive charges are called holes. Holes are quasiparticles because they merely describe the movement of electrons from one atom to another.

2.2.3 Semiconductor statistics

The availability of free carriers in the conduction band in semiconductors is a key factor in the operation of optoelectronic devices [19]. Free carriers can be introduced by doping if

thermal excitations can not create a sufficient amount of free carriers (which is in practice always the case). Doping means introducing elements having a higher or lower valence electron number. Higher electron number elements as dopants contribute conduction electrons and are called donors, and the lover valence number elements contribute holes and are called acceptors. Donor doped semiconductors are called n-type, and acceptor doped semiconductors are called p-type.

Note that even with doping, the amount of free carriers created with doping is temperature-dependent. In room temperature (300 K) it is safe to assume that all dopant atoms are ionized and the intrinsic thermal excited carriers are so few, that only carriers from doping need to be considered. Yet, when considering significantly lower or higher temperatures this is not the case. At lower temperature not all the dopant atoms are ionized, and at higher temperatures the thermally excited intrinsic carrier concentration becomes significant [17].

With classical (bosonic) particles, the energy distribution used is the Boltzmann distribution [19]. Carriers in semiconductors are fermions. Fermions differ from classical particles in that they are indistinguishable from another and they obey the Pauli exclusion principle. For fermions the Fermi-Dirac distribution must be used [17]. The distribution for electrons is

$$f(E) = \frac{1}{1 + e^{\frac{E-E_F}{k_B T}}}. \tag{10}$$

For holes the Fermi-Dirac distribution is one minus $f(E)$. This distribution tells the probability that a state with energy E is occupied. In the formula E is the energy of the state, E_F is the Fermi energy, k_B is the Boltzmann constant and T is temperature. The Fermi energy gives the highest occupied electron energy level. In metals the Fermi energy falls inside an allowed band, and in semiconductors inside the forbidden band. To be exact, the Fermi energy in the Fermi-Dirac distribution should be replaced with the chemical potential of the electrons, but it is traditional to use the form shown above [17]. At 0 K the chemical potential has the value of the "Fermi energy" that people commonly refer to. If $E - E_F \gg k_B T$ then the Boltzmann distribution approximates the Fermi-Dirac distribution well.

Electrons seek to minimize their energy and, therefore, the electrons occupy states with the smallest energy available. Because electrons are fermions no two electrons can have the same quantum state. Therefore, the more electrons there are in the solid, the higher states they must occupy. This means that with increasing donor doping the Fermi energy must rise. The surplus electrons must occupy the higher energy states. In semiconductors the Fermi energy lies in the forbidden zone, this means that increasing n-doping brings the Fermi energy closer to the conduction band edge. The opposite holds for p-doping.

At 0 K the Fermi-Dirac distribution is a Heaviside step function and states up to the Fermi energy are filled. The probability of electrons having higher energy than the Fermi energy is zero. This makes conduction impossible. This can be seen from the function $f(E)$ in

Eq. 10, which at 0 K is a step function. At higher temperatures electrons can occupy higher energies, and conduction is possible.

In addition to the Fermi-Dirac distribution, the density of states is required to know how many electrons lie at a specific energy. Density of states tells how many states lie between E and $E + dE$. This, with the Fermi-Dirac distribution completely specifies how many carriers there are in the conduction band. The density of states for different systems can be found in textbooks like Ref. 17.

2.3 The p-n junction

A p-n junction is an interface made from similar or dissimilar semiconductors having opposite doping. The perhaps most simple p-n junction imaginable is a silicon slab, where the other side is doped with acceptor atoms and the other side is doped with donor atoms.

When a p-n junction is formed there is an exchange of majority carriers. Some electrons diffuse from the n-side to the p-side to become minority carriers and vice versa. This leaves an area of ionized donors in the n-side and ionized acceptor in the p-side. The area is called the depletion zone. The depletion zone forms an electric field, see Fig. 3. The corresponding potential across the junction is called the built-in potential, and is an important quantity in the study of p-n junctions [19]. In the Schottky model the spatial charge is considered to be step-like (Fig. 3) and it enables a simple visualization of the electric field in the p-n junction. The junction reaches an equilibrium so that the diffusion current is exactly compensated by the drift current that is caused by the built-in field [19]. In equilibrium the drift current is constituted by the movement of minority carriers. As the junction allows electrons to exchange heat the junction reaches a thermal equilibrium. Therefore, the Fermi levels of the p- and n-sides shift to a common level. The thermal equilibrium causes also the bands to bend to form a common band through the depletion zone, see the uppermost picture in Fig. 4. The behaviour and the rectifying property of the p-n junction is presented in Fig. 4.

The fraction of conduction electrons that can overcome the potential barrier of the junction is determined by the Boltzmann factor

$$e^{-q\frac{V_{bi}-V_a}{k_BT}} \tag{11}$$

where q is the unit charge, V_{bi} is the built in potential and V_a is the applied bias voltage [17]. This applies also for holes in the p-side. The Boltzmann factor describes the rectifying property of the junction. When there is no bias few electrons can overcome the barrier. When the junction is under forward biased many electrons overcome the barrier and current is large. Reverse bias causes a neglible amount of current to flow. A more detailed analysis can be found in Ref. 17.

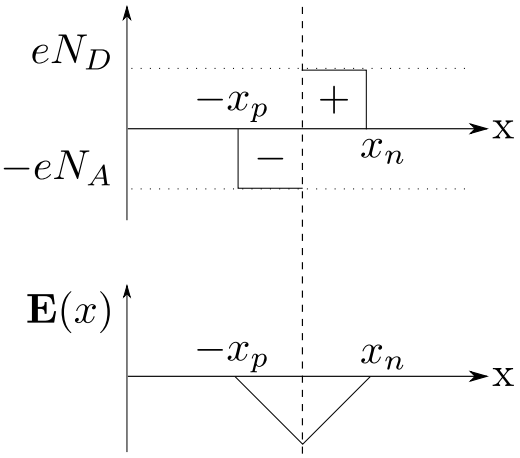

Figure 3: Spatial charge and the electric field in the Schottky model.

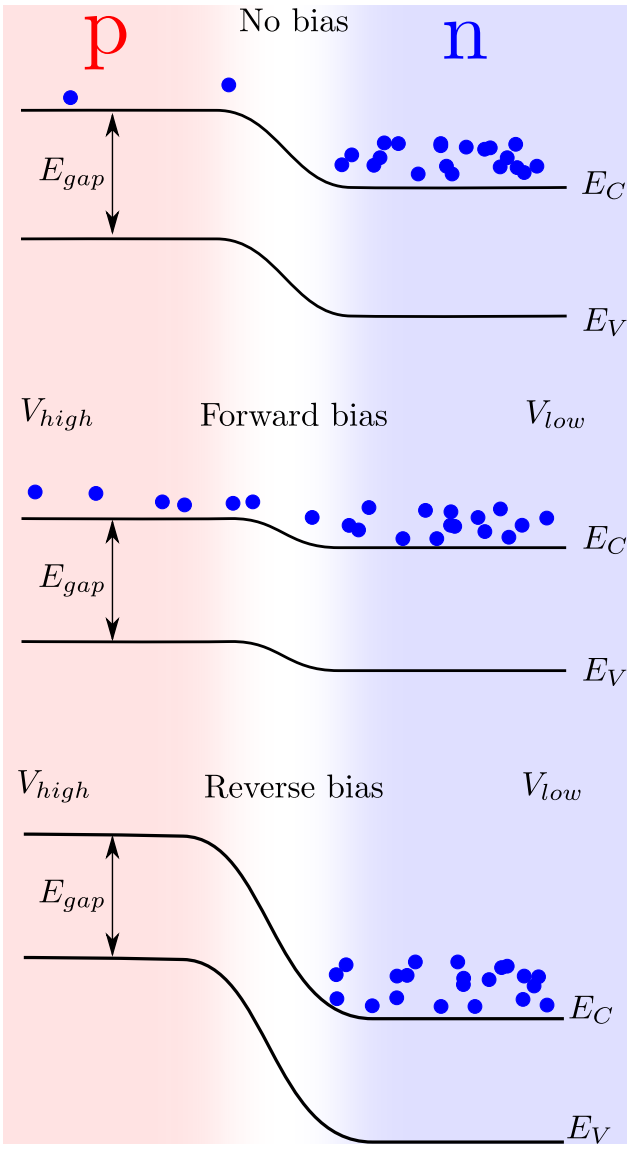

Figure 4: Movement of the conduction electrons (in blue) in a p-n junction. For clarity the movement of the holes is omitted. The topmost case portrays the p-n junction with no bias. In the middle the junction is under forward bias, and at the bottom the junction is under reverse bias. The *p*- and *n*-sides of the junction are colored red and blue, respectively.

2.4 Single quantum well LEDs

LEDs were discovered in 1904, and they are today often made from III-V semiconductors [19]. Lamps, display elements, sensors and even computer screens are some of the many of uses of the LED.

QWs are used in LEDs to produce nearly monochromatic light. QWs are usually made of compound semiconductors by inserting a lower band gap material into the bulk crystal, for example by introducing InGaN to GaN. The abrupt drop in the gap energy causes electrons in the conduction band, and holes in the valence band, to see potential walls on both sides.

QWs are 1D-potential wells, see Fig. 5. The energy states of an electron in an infinite 1D-potential box are obtained from solving the Schrödinger equation. The energies are

$$E_n = \frac{\hbar^2}{2m^*}\left(\frac{n\pi}{L}\right)^2, \; n = 1, 2, 3... \tag{12}$$

where L is the well width [19]. This means, that when the electrons recombine in a QW, the energies of the emitted photons are well defined. In the QW the electrons quickly thermalize (lose energy to phonons) to the lowest energy level of the QW (Fig. 5) [19]. When these electrons recombine nearly monochromatic light is produced. Thus, controlling the amount of lower band gap material in the QW (for example InGaN in GaN) the color of the emitted light can be controlled.

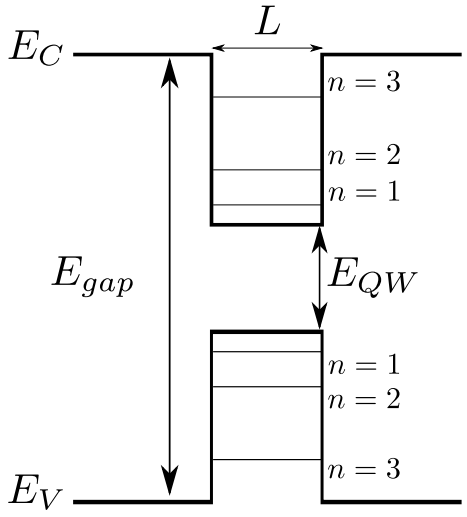

Figure 5: Energy levels of a QW shown schematically.

In an operating LED a forward biased p-n junction injects electrons into the p-side, to an empty conduction band. After a mean lifetime they recombine with holes in the valence band, releasing their excess energy as photons [19]. This process is known as spontaneous emission. This process in the case of a single quantum well (SQW) LED is presented in Fig. 6. Most of the electrons get trapped in the QW, and recombine producing light at the

wanted wavelength. Some electrons may end up to the p-side, where they recombine to produce a higher energy photon. Note that due to the applied voltage, the QW profile is not flat.

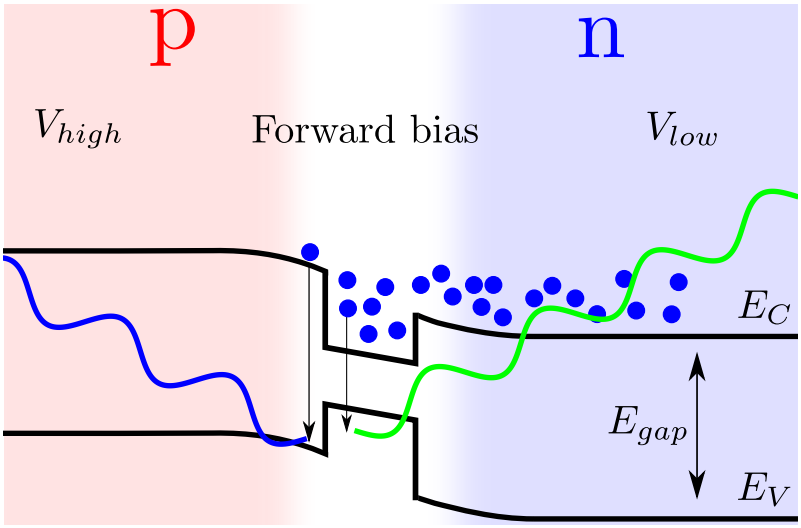

Figure 6: Operation of a SQW-LED under forward bias. For clarity only electrons are drawn. The wavy lines represents emitted photons from the recombinations.

3 Gallium nitride

The relevant compound semiconductor material for this thesis is gallium nitride (GaN). In this chapter we discuss its physical and optical properties, and the fabrication method used to fabricate some of the structures studied in this thesis (the QWs). The properties of GaN can be best understood by placing them in the context of the most used semiconductor, silicon. Therefore, GaN is compared here often with silicon.

This chapter describes also the structure of the samples used in the experiments. Since the distance of the QW from the surface is a crucial parameter in plasmonic applications, the unprocessed wafers were studied with XRR to determine the position of the QW. These results are presented here.

3.1 Physical properties

One of the most useful properties of GaN is its direct band gap of 3.44 eV. This corresponds to a vacuum wavelength of 360 nm, making GaN extremely favourable material for producing UV and near-UV region optoelectronic devices, such as LEDs, laser diodes and detectors. S. Nakamura *et al.* are regarded as the first to produce a high-brightness GaN based LED [20]. Silicon based LEDs are not commercially available, because the indirect band gap of silicon makes emitting photons very unlikely, as was discussed in Chapter 2. The wavelength would also be in the infrared in silicon LEDs. Though silicon can not in practice be used as a LED material, it can be used as a substrate for GaN growth [2, 21]. GaN LEDs on silicon are commercially available.

The band gap of GaN is considerably larger than that of silicon, see Table 1. The large gap energy also enables GaN-based transistors to operate at much higher temperatures than traditional silicon-based transistors. This is because thermal excitation of carriers in GaN is more difficult. Under low doping GaN also exhibits higher electron mobility than intrisic silicon (Table 1). Higher breakdown voltage of GaN transistors (up to 100 V) is also due to the larger band gap [17]. For these reasons GaN is preferred over silicon in fast high-power devices, like high-power transistors [17].

Gallium nitride is a very hard material, and GaN crystallizes into the wurtzite structure, with a considerably smaller unit cell size than that of silicon. Monocrystalline silicon has the diamond crystal structure. Both have similar hardness. Thermal conductivity of GaN is rather similar to monocrystalline silicon. This can be useful if GaN components are incorporated into silicon systems.

Table 1: Properties of GaN and Silicon [22–24]

	GaN	Silicon
Band gap / eV	3,44	1,13
Unit cell width / Å	3,19	5,43
Electron mobility / cm^2/Vs	500-3000	1500
Refractive index	2,41	3,42
Thermal conductivity / $\frac{W}{mK}$	130	150
Knoop hardness / kg/mm^2	750	850
Melting point / K	1160	1687

It would be unwise to not talk about the negative aspects of GaN, which are mostly cost and manufacture related. Manufacturing GaN is very expensive. Bulk silicon is grown from polycrystalline electronic grade silicon in the Czochralski process [24]. GaN, however, is often epitaxially grown from gaseous precursors in the metalorganic vapour phase epitaxy (MOVPE) process [2, 25]. Though MOVPE is fast compared to some other III-V growing techniques [25], producing silicon with the Czochralski process is much faster.

Another area where silicon is better than GaN is the abundance of the raw material. Silicon can be made from ordinary silica sand [24], whereas gallium is found only in trace amounts in the Earth's crust [26]. This naturally increases the price of gallium. Economies of scale are also in silicon's favor. Over 90% of microelectronics is silicon based [24]. This conveys a huge price advantage, and because silicon is the market leader in semiconductors it is naturally the first choice in any application in which silicon is a suitable material.

Growing GaN is difficult. This is mostly due to the lack of a lattice matched growth substrate [2]. High quality GaN films can be grown on sapphire using MOVPE (see Chapter 3.2), with the help of several buffer layers. The lattice mismatch between sapphire and GaN is partly mitigated this way, but the dislocation density (the amount of dislocations per unit area) is still large [2]. The quality of the GaN crystal would be enhanced, if a lattice matched substrate would become available.

GaN is famous for its ability to endure severe lattice mismatch. GaN grown on sapphire endures a lattice mismatch of 13% [2]. This high mismatch causes many dislocations in the crystal. The dislocation density is so large, that it is surprising that GaN based optoelectronic components work at all [2]! For example in conventional LEDs high dislocation density is associated with high nonradiative recombination.

The refractive index of GaN is high in the visible region of the electromagnetic spectrum, at an average of $n = 2,41$. In this thesis the visible region is considered to be from 400 nm

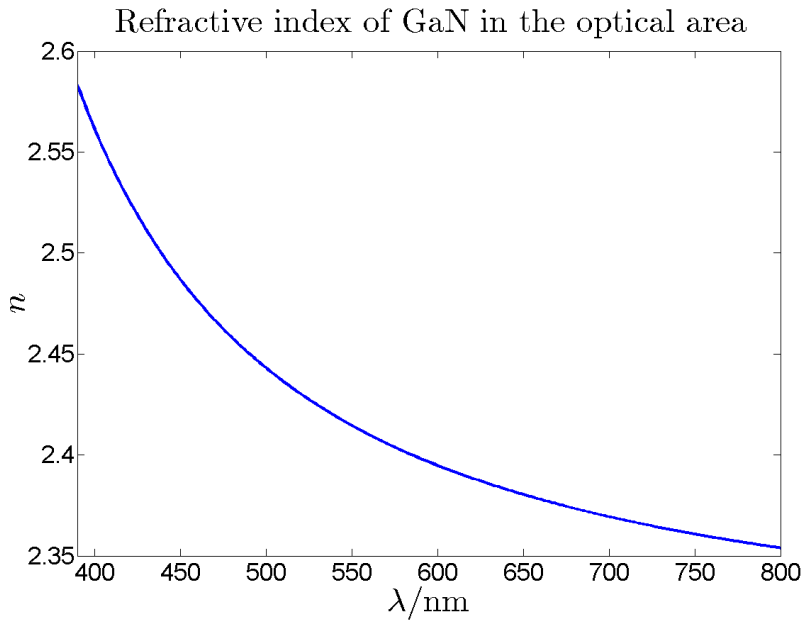

Figure 7: Refractive index of gallium nitride as a function wavelength. The curve was calculated with Eq. 20. The refractive index has no imaginary part in this wavelength range.

to 800 nm. In Fig. 7 the refractive index of GaN is calculated as a function of wavelength. The refractive index is $n = 2,47$ at the wavelength $\lambda = 470\,\text{nm}$ mostly used in this work. The high refractive index of GaN is a formidable challenge when considering LED applications, because this limits how much light can be emitted.

The amount of light emitted by a point source embedded in GaN can be estimated by calculating the angle to which emitted light does not experience total internal reflection at GaN-air interface. Assuming optically flat surface and neglecting the polarization properties of the transmission, the radiation efficiency is given by

$$\eta = \sin^2\left(\frac{1}{2}\arcsin\frac{n_{air}}{n_{GaN}}\right). \tag{13}$$

The radiation efficiency from the GaN/air interface at $\lambda = 470\,\text{nm}$ is $\eta = 4,3\%$. Because the QW can be regarded as a continuum of point sources, this is also the radiation efficiency for the QW. Improvements in the radiation efficiency are clearly needed. Methods have been developed to alleviate this problem [4–7, 9–12]. The method studied in this work is to couple part of the otherwise lost light into radiative modes through surface plasmon modes.

3.2 Metalorganic vapour phase epitaxy (MOVPE)

MOVPE is a mature and common fabrication process for fabricating compound semiconductor structures. MOVPE is valued for its suitability for large scale production and it can produce optoelectronic materials for a wide wavelength range, from UV to infrared.

Lasers, LEDs, compound semiconductor solar cells and ultra-high-speed transistors are just some of the applications made possible by using MOVPE processes [25]. Very abrupt interfaces can be made with MOVPE, and the crystal or doping profiles are indeed almost step-like.

In MOVPE organometallic reactant gases, such as trimethylgallium and trimethylindium, flow to a heated susceptor, where an appropriate seed crystal lies. The seed crystal is usually a crystalline wafer. A pyrolysis reaction occurs at the seed [25]. In pyrolysis an organic molecule is heated in an oxygen-free atmosphere, causing some of the molecular bonds to break, resulting in multiple reaction products. In the case of GaN, trimethylgallium can be used as the gallium source and ammonia as the nitrogen source. The gallium and nitrogen provided by the pyrolysis reaction bond with the seed crystal. This forms a new layer of GaN, where another pyrolysis can happen. The crystal grows this way. The seed crystal need not be the same material as the grown crystal, similar crystal structure and roughly similar unit cell are needed though. Mismatch between the lattice constants can cause strain on the wafer [19]. Too much stress at the interface can also even crack the wafer.

The other common method for producing compound semiconductor devices and crystals is molecular beam epitaxy. Although the quality of devices produced with molecular beam epitaxy is similar to MOVPE, the latter is the dominant manufacturing technology in compound semiconductors due to its scalability, versatility and the fact that MOVPE does not require slow and expensive vacuum technology [25].

3.3 GaN structures studied in this thesis

The GaN structures used in this work were grown with MOVPE to single-side polished sapphire substrates, by M.Sc. Olli Svensk. The wafers all have a SQW with a band gap energy corresponding to 2.64 eV or $\lambda = 470$ nm and the well thicknesses are 2.5 nm. Stoichiometrically the SQW consists of $In_{0,15}Ga_{0,85}N$. A schematic of the used wafers is presented in Fig. 8. The topmost GaN layer is called the capping layer. The thickness of the capping layer is a critical parameter in the plasmonic coupling, to be explained in detail in Chapter 6. The capping layer widths were measured using XRR, see Table 2. The errors in Table 2 are given by the diffractometer software. A detailed description of XRR is given in Parrat's pioneering paper [27]. Because the the uniformity of the MOVPE process the capping layer thicknesses are presumed to be equal between the wafers of each series. Each series is grown in the same chamber during the same growth run, and consists of three wafers. Therefore only one wafer from each series was measured. The 150 nm capping layer could not be measured using XRR. The diffractometer does not give accurate XRR measurements to this depth. Luckily the 150 nm samples are for comparison and according to the MOVPE operator, the capping layer width is at least 100 nm. This is enough distance from the QW to seriously attenuate the coupling. This is discussed more in Chapter 6.

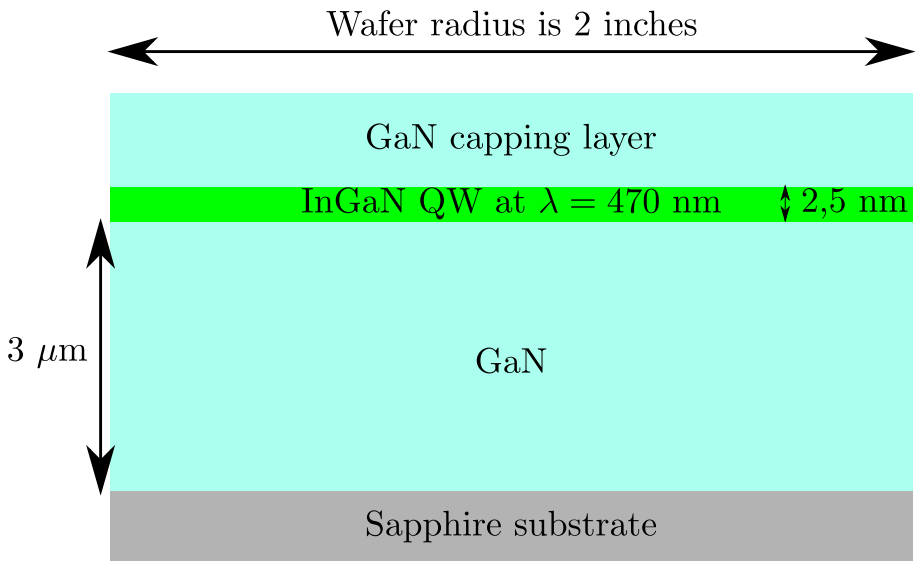

Figure 8: Structure of the wafers used in this work. The capping layer thickness is depending on the sample either 5 nm, 10 nm or over 100 nm.

Table 2: Thicknesses of the capping layers measured by XRR.

Desired	Measured
5 nm	5,4 nm ± 0,2 nm
10 nm	10,2 nm ± 0,4 nm

4 Surface plasmons and plasmonic coupling

This section first explains what plasmons and surface plasmons (SPs) are. The properties of SPs are discussed in detail, specifically the fields and the dispersion relation of SPs. The suitability of GaN and silver for sustaining SP at the energy of interest are then reviewed. Then the SP dispersion relation for the GaN/Ag interface is calculated. Finally, plasmonic coupling mechanisms are presented and how these can be used in the case of the InGaN QW.

4.1 The nature of plasmons

Plasmons can be viewed as collective oscillations of the free electrons in conductors. There are three types of plasmons commonly discussed: the volume plasmon, the SP and the localized SP [28].

The volume plasmon is a longitudinal oscillation of free electrons, and as such it can not couple to radiation. Volume plasmons can be excited via an electron beam [29]. They can be thought of as density oscillations of the free electrons.

SPs are electromagnetic waves propagating along a metal-dielectric interface, confined in the perpendicular direction [28]. The fields in SPs are coupled to the surface charge oscillations of a free electron gas (usually in metal), see Figs. 10 and 11. The most important phenomenon in this thesis, is the coupling of energy to SPs. Due to the transverse nature of the SP it can be coupled to radiation, but coupling requires momentum conservation considerations. These are discussed in section 4.3.

Metallic nanoparticles and nanostructures are the third branch where localized SPs exist. Plasmonic resonances in metallic nanoparticles give them unique size, ordering and dielectric environment dependent optical properties [8, 28]. These properties have been historically used to stain glass. Metallic nanoparticles were diffused into the glass to colour it, where they acted as wavelength dependent scatterers. The lively colours of many cathedral windows are due to the nanoparticle plasmon resonance, see Fig. 9.

4.2 Surface plasmons

To understand SPs it is necessary to understand the physics behind them. That physics is discussed here. As the SP can exist only in certain material interfaces, those considerations are also discussed here. The plasmonic properties of the materials of interest in this work, GaN and silver, are investigated thoroughly.

Assuming time-harmonic solutions to Maxwell's equations, the Helmholtz equation for **E** and **H** gives:

Figure 9: Cathedral window. Image from Ref. 30 .

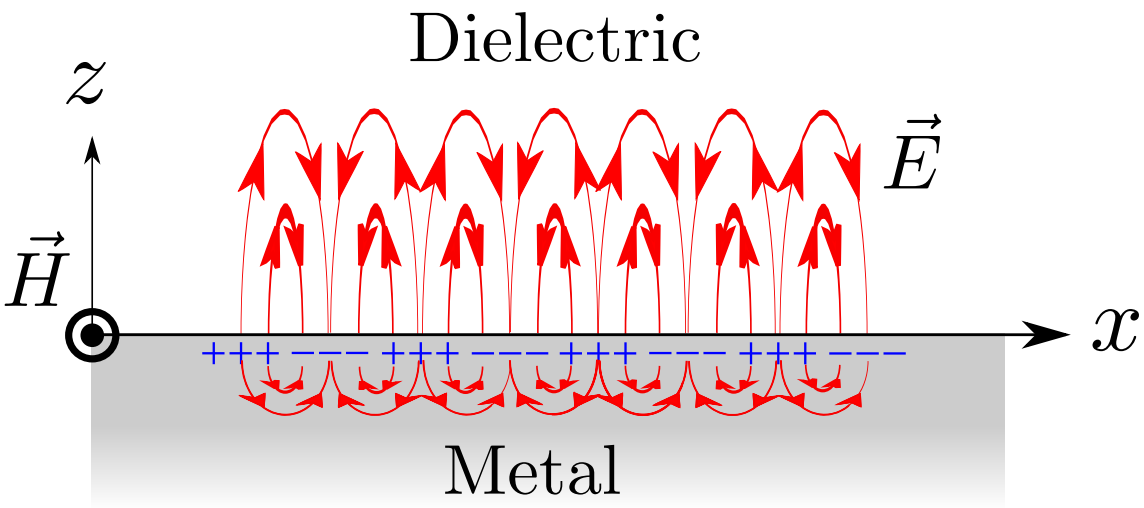

Figure 10: SP electric field in red and surface charge density oscillations in blue. Note that, because of the negative permittivity of the metal, the electric field dies faster.

$$\nabla^2\mathbf{E} + {k_0}^2\epsilon\mathbf{E} = 0 \qquad (14)$$
$$\nabla^2\mathbf{H} + {k_0}^2\epsilon\mathbf{H} = 0$$

The SP arises as a solution to these equations at the metal-dielectric interface. The derivation is straightforward (see, *e.g.*, Ref. 28). Note, that only transverse magnetic (TM) solutions exist. TM means here that the magnetic field has no component perpendicular to the interface. The TM solution to the wave equation is

$$\begin{aligned} H_y(z) &= A_D e^{i\beta x} e^{-k_D z} \qquad (15)\\ E_x(z) &= iA_D \frac{1}{\omega\epsilon_0\epsilon_D} e^{i\beta x} e^{-k_D z}\\ E_z(z) &= -A_D \frac{\beta}{\omega\epsilon_0\epsilon_D} e^{i\beta x} e^{-k_D z}\\ k_D^2 &= \beta^2 - k_0^2\epsilon_D \end{aligned}$$

for $z > 0$ (*cf.* Fig. 10) and

$$\begin{aligned} H_y(z) &= A_M e^{i\beta x} e^{k_M z} \qquad (16)\\ E_x(z) &= -iA_M \frac{1}{\omega\epsilon_0\epsilon_M} e^{i\beta x} e^{k_M z}\\ E_z(z) &= -A_M \frac{\beta}{\omega\epsilon_0\epsilon_M} e^{i\beta x} e^{k_M z}\\ k_M^2 &= \beta^2 - k_0^2\epsilon_M \end{aligned}$$

for $z < 0$ [28]. The subscripts M and D correspond to the metal and dielectric side, respectively. The z-component of the wave vector is abbreviated $\mathbf{k}_{z,i}$=$\mathbf{k}_i$, A is the amplitude of the magnetic field and $\beta = k_x$ is the propagation constant. The fourth equation is a necessary condition that arises because H_y has to fullfil the Helmholtz equation in both domains. Fig. 12 represents the oscillation of the SP electric field as a function of x when time is constant.

According to Eq. 15 and 16, for the wave to be confined to the interface, $\mathrm{Re}\,(\epsilon_M) < 0$ and $\mathrm{Re}\,(\epsilon_D) > 0$. This is because when moving away from the interface the exponential depending on z must converge to zero in both media. If the fields are not defined evanescent, they are not confined in any meaningful way. Opposite signs for the real parts of the permittivities of the materials are therefore needed for SPs to exist [28]. Typically this is accomplished with a metal and a dielectric. Any two materials fulfilling the permittivity condition are sufficient. The frequency dependency of permittivity complicates the analysis. This is discussed later on.

Using the well-known boundary conditions for the perpendicular tangential parts of the electromagnetic fields on the interface, a new condition for the propagation constant β is obtained [28]:

$$\beta = k_0 \sqrt{\frac{\epsilon_M\epsilon_D}{\epsilon_M + \epsilon_D}}\,. \qquad (17)$$

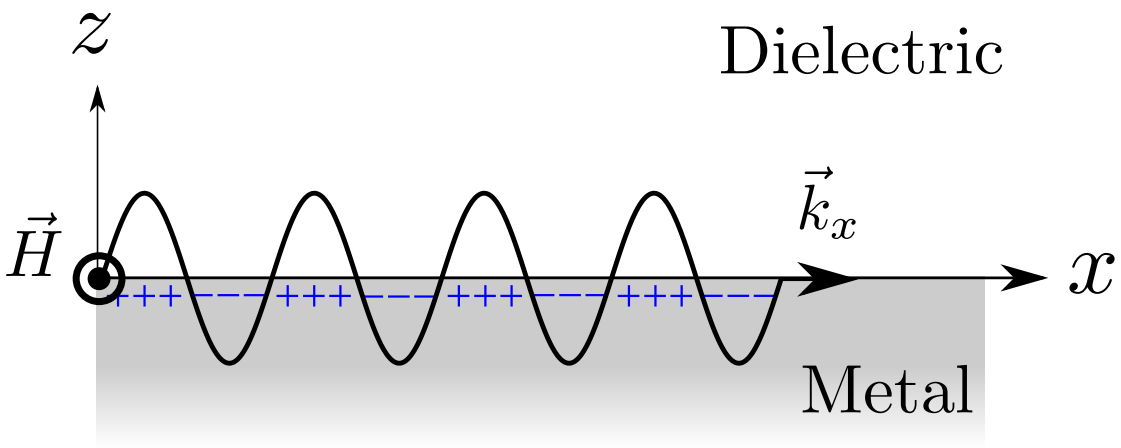

Figure 11: Surface charge density oscillation coupling to a SP is represented by the charge signs. The black line is the magnitude of the x-component of the electric field. The charge oscillation is in the xy-plane.

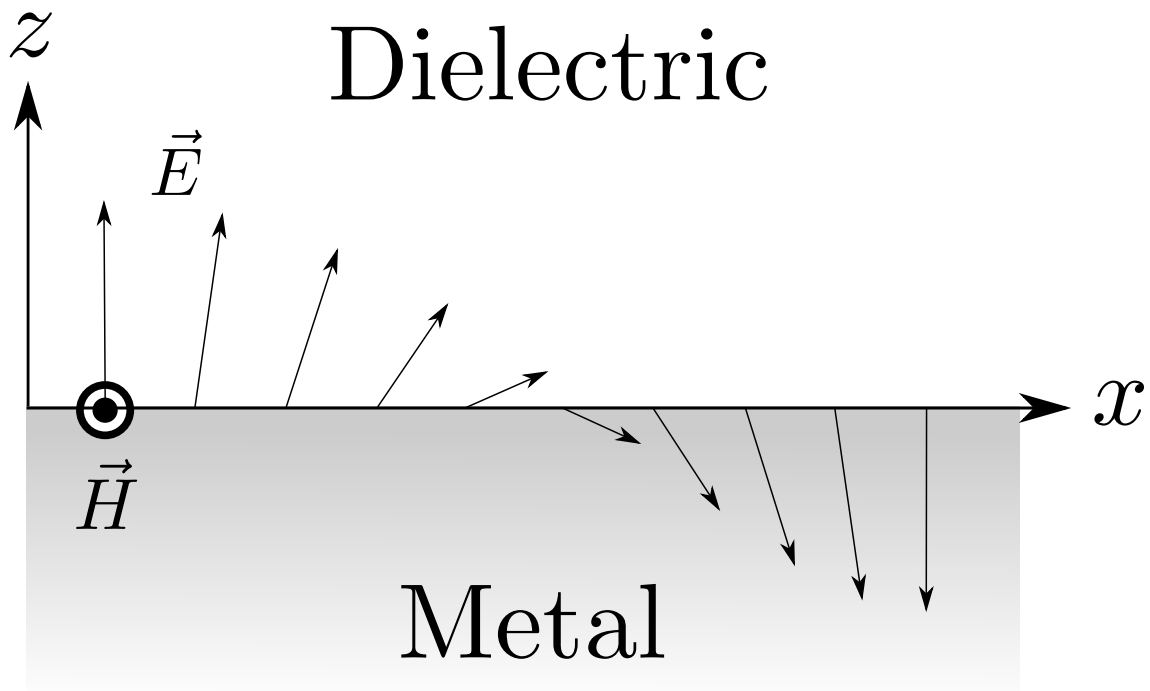

Figure 12: Oscillation of the SP electric field as a function of x when time is constant. The electric field changes in the x- and z-directions as x changes. The magnetic field oscillates only in the y-direction.

This is the dispersion relation for the SP.

The Drude model, named after the German physicist Paul Drude, describes the metal as a free electron gas that moves in the presence of the positive cores [28]. The model pictures the electrons oscillating due to an external electric field, and that the electron motion is damped by collisions, at a frequency γ [28]. Solving the associated classical second-order differential equation for the electron motion and remembering the connection between polarization and the electric field one arrives at

$$\epsilon_M(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}, \tag{18}$$

where γ is the collision frequency and ω is the angular frequency of the electric field associated with the incoming light [28]. By inserting this into Eq. 17, we see that when wave vector is large Eq. 17 approaches

$$\omega_{SP} = \frac{\omega_p}{\sqrt{1+\epsilon_D}}, \tag{19}$$

where ω_{SP} is called the SP frequency.

Eq. 18 fails to account for the increase in the imaginary part of ϵ_M due to interband transitions that occur at sufficiently high energies [28]. The interband transitions of silver start at $E = 3.98\,\text{eV}$ ($\lambda = 312\,\text{nm}$) [31]. Therefore, interband transitions need not to be taken into account in the visible region for silver. Gold, on the other hand, has the interband transition energy corresponding to $\lambda = 550\,\text{nm}$ [28]. This, and the complex permittivity of gold predicted by the Drude model at near-infrared, explain why gold is yellow [28].

The plasma frequency and collision frequency of silver are given by Ordal *et. al* [32] to be $\omega_{p,Ag} = 1.38 \times 10^{16}\,1/\text{s}$ (139 nm) and $\gamma_{p,Ag} = 4.35 \times 10^{12}\,\text{Hz}$ [32], respectively. The energy corresponding to the plasma frequency is well outside the energy of interest. Therefore, silver retains its metallic character in the visible region. This is important, because the Drude model predicts that metals are transparent for frequencies greater than the plasma frequency. The measurements agree very well with the Drude model in the visible region [17]. If the frequency corresponding to the energy of interest would be greater than the plasma frequency, the GaN/Ag interface would be just a dielectric-dielectric interface, where SPs can not exist.

The properties and behaviour of silver can be calculated in the visible region. The refractive index (thus the permittivity) of GaN is also needed for calculating the SP resonances. For GaN the Sellmeir equation is

$$n^2(\lambda) = 3,60 + \frac{1,75\lambda^2\times10^{12}}{\lambda^2\times10^{12}-(0,256)^2} + \frac{4,1\lambda^2\times10^{12}}{\lambda^2\times10^{12}-(17,86)^2}, \tag{20}$$

and it predicts the refractive index as a function of wavelength [23].

Because the applicability of the Drude model has been verified with the materials used in this thesis in the visible region, the dispersion relation for the SP can be calculated.

The dispersion relation in the visible region is displayed in Fig. 13. It is customary to present the dispersion relation with the frequency on the y-axis, although usually people are interested in the β as a function of frequency. In Figs. 13 and 15 the frequency is expressed as the normalized angular frequency $\frac{\omega}{\omega_p}$ and the wave vector is also expressed in non-dimensional form $\left(\frac{Re\{\beta\}c_0}{\omega_p}\right)$. This convention is also used in Ref. 28. The visible region is between 0.17 and 0.34 of the normalized angular frequency.

The light line represents the maximum value of k_x the light can have in that medium without becoming evanescent. Because $k^2 = k_z^2 + k_\rho^2$, there is an upper limit for the k_x, before k_z turns imaginary, and emission becomes evanescencent. It can be seen that for a certain energy, the SP has always a larger k_x than photons at the same energy, see Fig. 13. This holds for vacuum, air, GaN and any other dielectric. This means that surface plasmons can not be directly excited with radiation, and vice versa, SP cannot directly radiate, because momentum must be conserved. This momentum mismatch is of key importance.

The light cone is a visualization the maximum k_ρ the radiating light can have. Here, and in literature [28], "inside the light cone" means "on the left side of the appropriate light line". A depiction of the light cone can be found in Fig. 14.

At higher energies the dispersion relation for β is more complicated, see Fig. 15. Lower part of the left picture contains the classic "text book" SP dispersion relation. The visible region is between 0.17 and 0.34 of the normalized angular frequency (marked in green). The upper part of the same picture has the area marked as "Radiates". There the mode is radiative, because the SP solution is inside the GaN light line the SP can couple to radiative modes. This area begins at the plasma frequency, where the silver becomes transparent. Radiation to silver below the plasma frequency is not possible.

With increasing energy the GaN light line cannot really be called a line anymore (Fig. 15). Fortunately the refractive index of GaN changes very little in the visible region (see Figs. 7 and 15) so the GaN light line is indeed almost a straight line. It is also interesting, that the component k_x in GaN becomes purely imaginary between 0.54 and 0.66 of the normalized angular frequency (Fig. 15). This means that GaN is opaque in this region. The opaque region is also well outside the energy of interest.

Between the classic area and the plasma frequency is the area where quasi-bound modes exist [28]. There SPs are evanescent in the x-direction and propagate in the z-direction. What is surprising is the area marked "SP-like modes". There the β of the SP is on the right side of the GaN light line, so it cannot radiate into the gallium nitride. Therefore, it is confined in the z-direction, and the propagation constant is real. This matches the properties of the normal SP, confinement in the z-direction and propagation in the x-direction. Because these solutions are so far outside the visible region they do not need to be considered experimentally in this work (see Chapter 6). However, theoretically they are somewhat interesting.

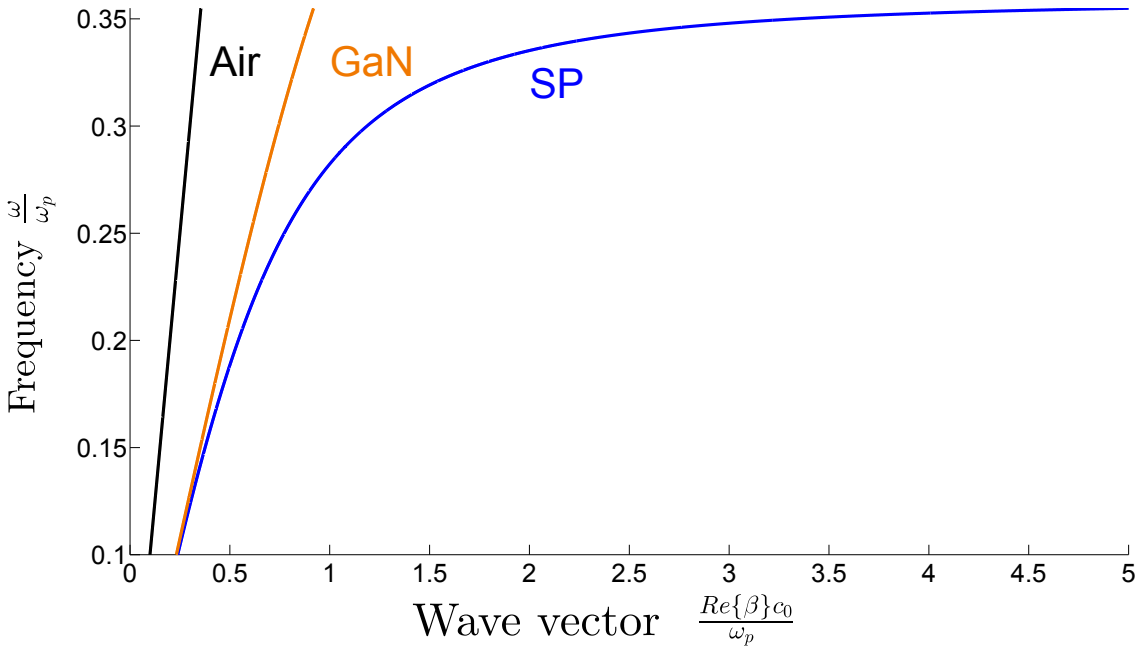

Figure 13: Dispersion relation of the SP on a GaN/Ag interface. This is the classical "text book" dispersion relation of the SP at low energies. The vacuum or air light line is in black and the light line in GaN is in orange.

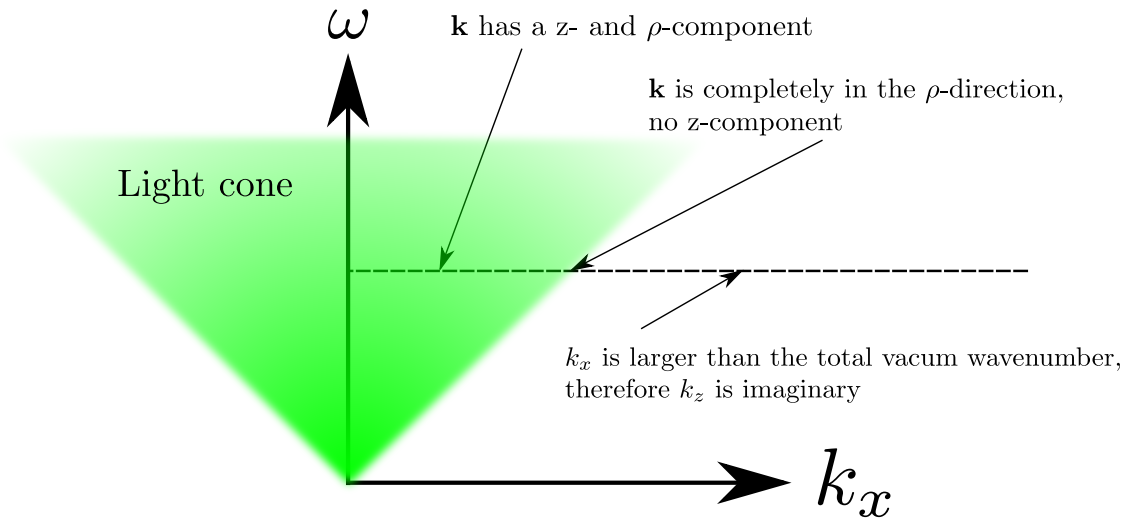

Figure 14: Schematic of the light cone is shown in the picture. The dashed line represents an arbitrary frequency for the light. Note that only real values of **k** are able to radiate, the modes with complex **k** are evanescent.

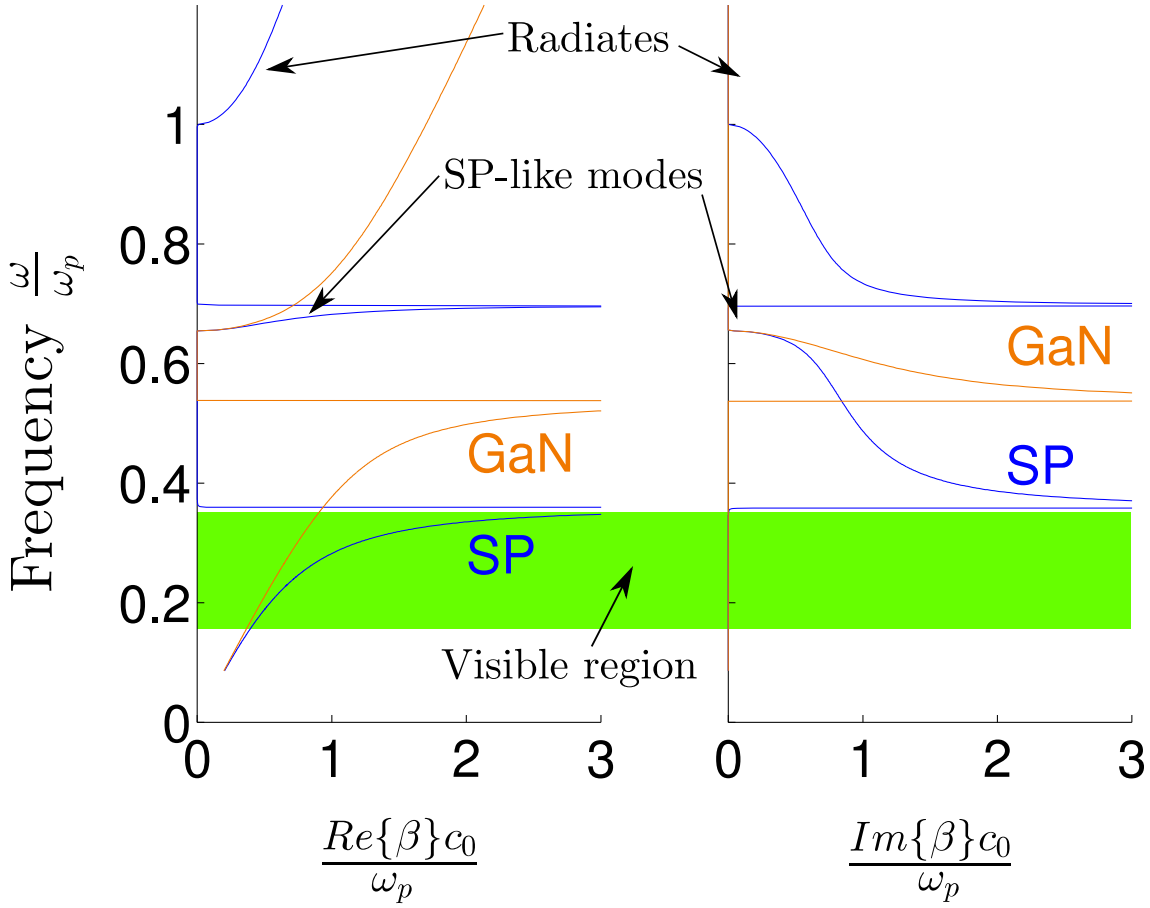

Figure 15: On the left: the real part of the SP dispersion relation and the GaN light line. On the right: the respective imaginary parts. Both have in the y-axis the normalized angular frequency. The increase of the imaginary part of the permittivity of silver at higher energies is not included in this calculation. In addition Eq. 20 might not be valid at such large energies. The areas pointed out in the figure are discussed in the text. The green area corresponds to the visible region.

It has been demonstrated that silver can support SPs, but is silver a good choice. Silver is suitable, because it has low losses in the visible region. This can be seen from the right-hand-side diagram in Fig. 15. In the visible region, coloured green, the imaginary part of k_x is neglible. This means that SPs experience low losses when propagating along the interface. In visible and near-infrared plasmonics gold and silver are commonly used [28], but usually silver is preferred due to the lower losses.

The structures fabricated and measured in this work have three layers: GaN, silver and air. This is called an insulator-metal-insulator structure (IMI) [28]. This means, that the SP dispersion relation Eq. (17) which was derived for a dielectric-metal interface is not strictly correct. IMI structures have slightly more complicated fields that are derived in Ref. 28. Both metal surfaces have a field similar to a single metal-dielectric interface SP. The field propagates along the surface and is evanescent in the z-direction. Inside the metal the field is a superposition of these two fields. The dispersion relation of IMI structures is

$$e^{-4k_M a} = \frac{\frac{k_M}{\epsilon_M} + \frac{k_{D,1}}{\epsilon_{D,1}}}{\frac{k_M}{\epsilon_M} - \frac{k_{D,1}}{\epsilon_{D,1}}} \frac{\frac{k_M}{\epsilon_M} + \frac{k_{D,2}}{\epsilon_{D,2}}}{\frac{k_M}{\epsilon_M} - \frac{k_{D,2}}{\epsilon_{D,2}}} \tag{21}$$

Where a is the thickness of the metal. For infinite metal thickness Eq. 21 reduces to two single interface SP [28]. The dispersion relation for IMI structures is qualitatively very similar to a metal-dielectric interface [28].

4.3 Coupling surface plasmons to light

It was discussed that SPs can not radiate from the metal-dielectric interface due to the momentum mismatch, which also means that SPs cannot be excited with light due to the same momentum mismatch. Provided that the momentum is somehow compensated, SPs can be excited with light and vice versa. In this section, several methods for SP coupling are discussed

The classical method of exciting SPs is prism coupling. The two classical prism coupling configurations are presented in Fig. 16. The incident light, in this case a laser beam, undergoes total internal reflection. This causes the light to become evanescent in the second material, because of the purely imaginary wavevector in the z-direction

$$\begin{aligned} \beta &= \sqrt{k_{prism}^2 + \mathrm{Im}(k_z)^2} \\ \beta &= \sqrt{\left(\frac{\omega}{c_0} \times n_{prism}\right)^2 + \mathrm{Im}(k_z)^2}. \end{aligned} \tag{22}$$

Thus, the propagation constant becomes larger than the total wavevector in vacuum, and this enables the light to excite SPs. Note, that momentum matching on the prism-metal interface can not be achieved, because the prism-metal interface is just another metal-dielectric interface for the incoming laser light. The SP is excited on the air-metal interface.

Though prism coupling is simple in theory, in practice it is more difficult to implement. In the Kretschmann-Raether configuration (Fig. 16, on the left) the metal layer needs to be extremely thin, tens of nanometers, because the field is evanescent. Due to the evanescence the metal layer is usually evaporated on the prism [28]. The Otto configuration (Fig. 16, on the right) is even harder to realize in a laboratory, because the air gap needs to be extremely precisely controlled [8].

Grating coupling can used to overcome the momentum mismatch [8, 28]. The interface is patterned with a grating, grooves or a series of holes [28]. Any periodic pattern of holes or grooves is a grating. The grating is usually milled into the metal, but the grating can also be made from a dielectric [28]. A one dimensional grating is shown in Fig. 17, with a period a. The grating adds a reciprocal-lattice vector to the propagating SP (or to the incoming light) [28]. This enables coupling to light (and from light) , when

$$\beta = k_{x,light} \pm n\frac{2\pi}{a}, n \in \mathbb{Z}, \tag{23}$$

is fullfilled [28]. This condition applies when the SP is propagating along the so called grating Bragg vector, which is perpendicular to the stripes (Fig. 17). The periodicity must continue over an extended region [28].

Because the radiated SP needs the periodicity of the grating to radiate, the SP has to propagate along the surface so that it at least has a small component along the grating Bragg vector (Fig. 17). Ref. 33 observed that 70% of the energy of a buried emitter, like a QW, can couple to SP modes [33]. This was a classical result, which should be remembered later when discussing non-classical coupling. Ref. 33 estimates that half of all SP propagation directions result in grating scattered emission (Eq. 23) and that for these directions the coupling efficiency of a buried emitter is 70% [33]. This would yield a total efficiency of roughly 25%. For a 2D grating the efficiency would be then 50%, according to Ref. 33. A 2D grating has 2 orthogonal Bragg vectors, see Fig. 40 in the next chapter. These estimations were for the single metal-dielectric interface. For a three layer structure, like the ones that were fabricated for this work, SP emission efficiencies of roughly 50% have been measured [33].

The classical SP coupling seems rather efficient by itself if a 25% total coupling efficiency can be attained. In addition to the classical coupling there is another mechanism to couple the QW emission to SPs and this to light. The principle is presented by I. Gontijo *et al.* [13]. They studied the coupling of an InGaN QW to SPs on a planar silver/GaN interface. An InGaN QW was 12 nm below a 8 nm silver layer. The photoluminescence from this silver covered area was compared with an area not covered with silver. A spectrally sharp photoluminescence drop, by a factor of approximately 55, was observed [13]. The drop is due to the electron-hole energy being rapidly transferred to the SP mode [13]. The mechanism observed is strong quantum electrodynamical coupling. Note that no light travels between the QW and the silver/GaN interface. The spatial overlap of the SP and the electron-hole pair electric fields allows this non-classic phenomenon to occur. The precise character of the strong quantum electrodynamical coupling is outside of this work. Luckily the coupling obeys Fermi's golden rule [13]. The recombination rate directly to

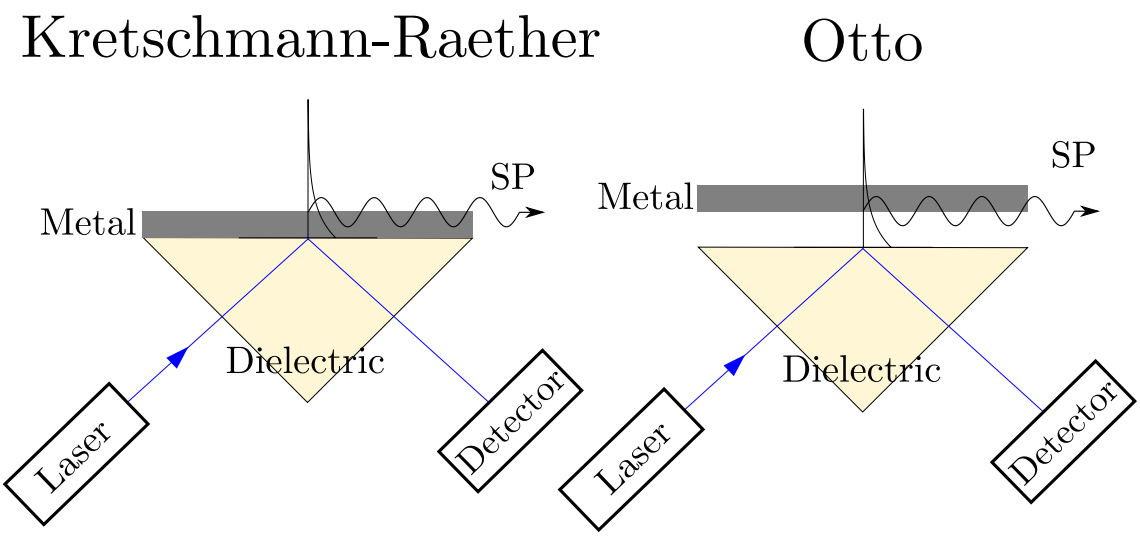

Figure 16: The classical ways of coupling light into SPs, the Kretscmann-Raether configuration on the left and the Otto configuration to the right. In both confugurations the light is totally reflected from the dielectric air interface. The evanescent field produced can couple to the SP. The evanescent field is depicted in the picture by the exponential curve that drops as z-increases. Picture inspired by Ref. 8.

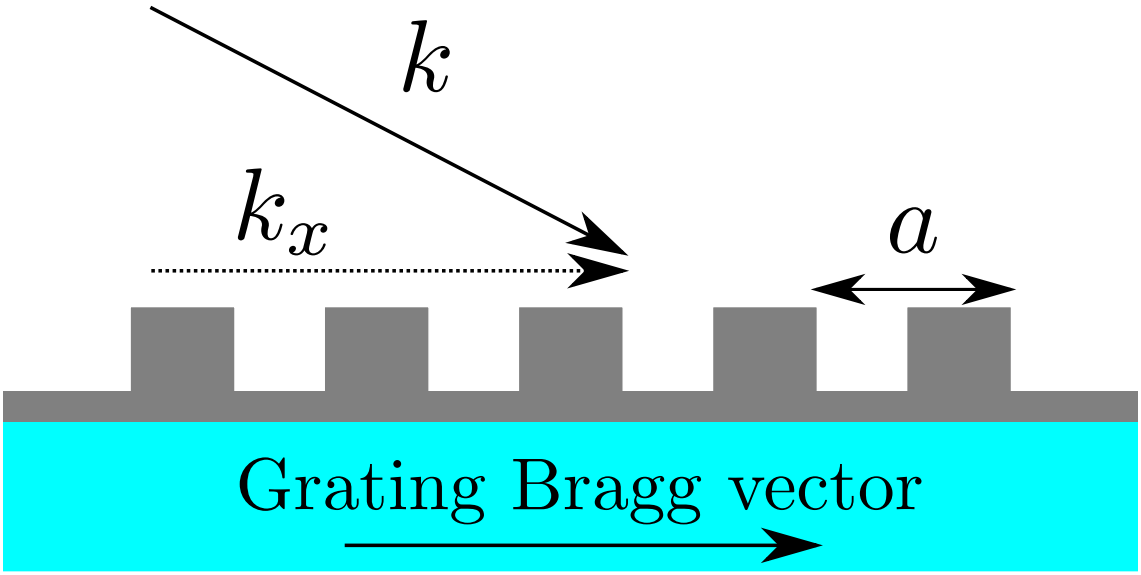

Figure 17: Momentum-matching of light to SP modes using grating coupling. The period of the grating is a, k is the wavevector of the incoming light and k_x is the wavevector component in the direction of the grating Bragg vector.

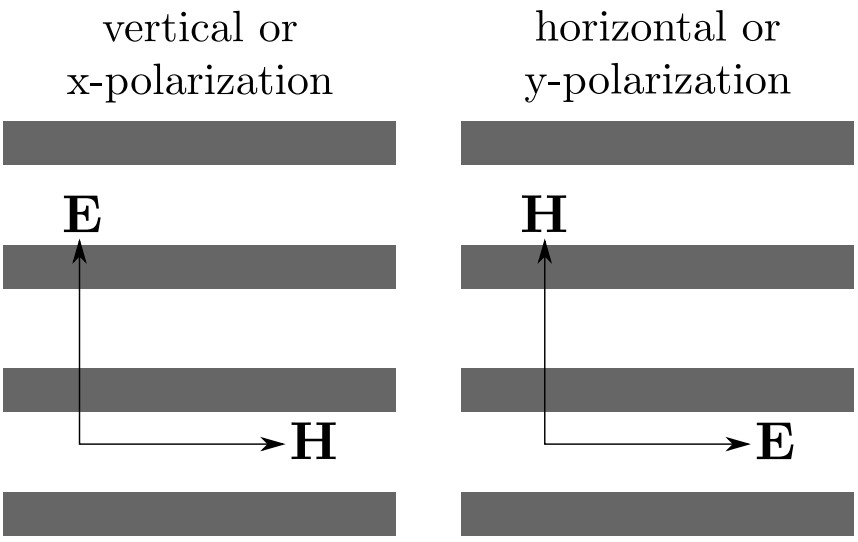

Figure 18: Polarization convention for the SP and emitted light is pictured. The gray bars represent the silver grating.

SP modes can be thus estimated to be

$$\Gamma_{SP} = \frac{2\pi}{\hbar} <\mathbf{p} \cdot \mathbf{E}(a)>^2 \rho(\hbar\omega) \tag{24}$$

where $\mathbf{p}$ is the dipole moment of the electron-hole pair, $\mathbf{E}(a)$ is the electric field of the plasmon at the QW depth (capping layer width) and $\rho(\hbar\omega)$ is the plasmon density of states at the bandgap energy [13]. It can immediately be seen that because the SP is an evanescent mode, the recombination in this fashion quickly dies off when the QW is positioned further and further away from the interface. If the QW can be positioned near to a silver layer, and if efficient antenna nanostructures can be incorporated in the silver film, the energy of the QW can be efficiently transferred to radiative modes outside the semiconductor [13].

Because the SPs exist only in the TM-polarization, and the SP has to have a component collinear with the grating Bragg vector to radiate [33], the SPs that are coupled to radiation are polarized. This is due to simple conservation of momentum. When the magnetic field of the SP is perpendicular to the grating Bragg vector, this will be called x-polarized or vertical polarization. Similarly, when the magnetic field is collinear with the grating Bragg vector, the SP is y-polarized or horizontally polarized. The vertical and horizontal naming come from the alignment of the samples in the measurements, see Chapter 6. An illustration of the polarizations is shown in Fig. 18. To summarize, the horizontal polarization is not likely to radiate. This point is crucial in the SP coupling experiments presented in Chapter 6.

5 Nanofabrication of plasmonic gratings

The fabrication process of large area subwavelength plasmonic silver gratings both on silicon and on GaN using azo-polymer etch masks is explained in detail in this section. First, the general properties of azo-polymers are discussed, then the surface relief grating (SRG) formation on azo-polymers is explained and finally the manufacturing process is presented. The fabrication of two-dimensional (2D) gratings is also presented.

5.1 Azo-polymers

Polymers having different azobenzene groups are called azo-polymers. The elongated form of the azobenzene is seen in Fig. 19. The first use of this molecule was incorporating it to different dyes to produce vivid colors [34]. This can be well seen even in dilute solutions of azo-polymers, which can have extremely bright red and orange colors. For example the 2 wt% (weight percent) solution of poly(Disperse Red 1 acrylate), abbreviated pDR1a, in 1,2-dichloroethane has a striking blood red color, see Fig. 20. The 1,2-dichloroethane is a powerful solvent that is used to dissolve the azo-polymer, and to spin coat the azo-polymer. The chemical structure of this azo-polymer with the chromophore (the part responsible for color) in the *trans* isomer form can be seen in Fig. 21.

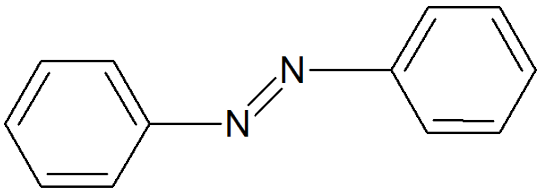

Figure 19: Azobenzene molecule.

Figure 20: Photograph of a 2 wt% solution of pDR1a in 1,2-dichloroethane.

The azobenzene or azo group can undergo *trans-cis* (or *cis-trans*) isomerization under UV or visible light [34], a process known as photoisomerization. The *trans* state has lower energy than the *cis* state, so eventually all *cis* isomers relax thermally back to the *trans*

form, *cf.* in Fig. 23 [34]. The thermal relaxation can take from seconds to hours, whereas the photochemical process is in the picosecond scale [34]. These three processes in azo-polymers were first observed in the solid state by Paik and Morawetz in 1972 [35].

Many different azo groups and azo chemicals can undergo rapid photoisomerization [34]. This property is retained even when the azo group is in a chromophore bonded to a side chain of a polymer as in pDR1a, see Figs. 21 and 22. These conformal changes in the polymer chains can affect the chemical and physical properties of the polymer [34].

The photoisomerization of azo groups is polarization-dependent. This is fairly straightforward, because the rate of photoisomerization depends on the probability of absorbing a photon. Furthermore, the chance of the molecule absorbing a photon this way is dependent on the cosine squared of the angle between the transition dipole moment of photoisomerization and the electric field of the excitation [34]. This can easily be understood via an analogy with a "normal" absorption dipole moment [8, 34]. The dipole moment in question with the polymer is the transition dipole moment of the chromophore [34]. Note that the *trans-cis* transition dipole moment and the *cis-trans* transition dipole moment of the chromophore need not be the same. When writing SRGs it is preferrable that the maxima of the transition dipole moments coincide at a specific wavelength. Care must be then taken to choose an azo chemical where these dipole moments are high at a specific wavelength. Irradiating the azo chemical with the suitable wavelength causes the chromophores to undergo repeated photoisomerizations until their transition dipole moments are perpendicular to the field. This can be used to make birefringent holographic gratings as first reported by Todorov *et al.* [36].

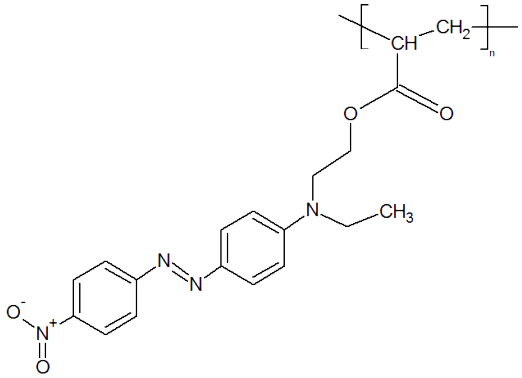

Figure 21: pDR1a with the azobenzene in the *trans* form.

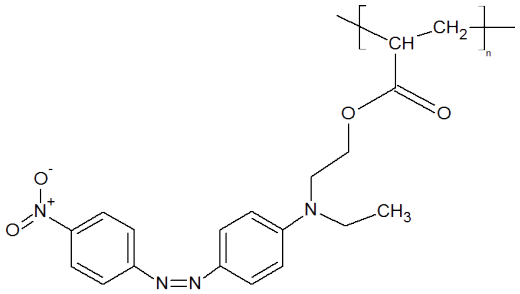

Figure 22: pDR1a with the azobenzene in the *cis* form.

5.1.1 Surface relief gratings

The azobenzenes have been studied for a long time, but a completely new line of research was discovered in 1995, when two papers regarding optically induced surface relief gratings in azo-polymers were simultaneously and independently published [37, 38]. Both

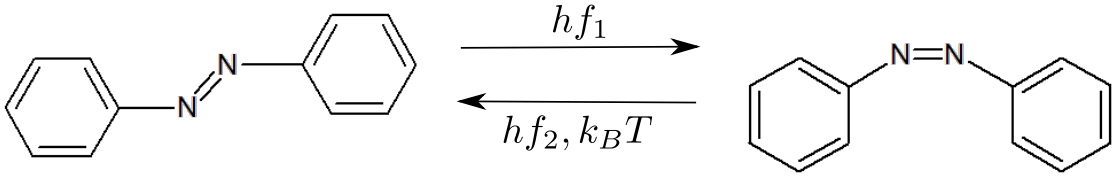

Figure 23: The azobenzene molecule undergoing *trans-cis* isomerization under the influence of a photon of frequency f_1 and undergoing *cis-trans* isomerization under the influence of a photon of frequency f_2 or heat. The *trans* isomer is on the left and the *cis* isomer is on the right.

groups had inscribed SRGs on thin (under one micron) azo-polymer films by exposing the film to interfering light. The year before there was a publication on this topic, but it was not in a peer reviewed science journal [39]. The findings were completely unexpected. Large diffraction efficiency from the irradiated surface was detected. Diffraction efficiency is a figure of merit for a grating. It is defined as the ratio of the power of the laser beam that hits the surface divided by the power in the first diffraction maxima. This was higher than was expected from previous experiments with photoinduced birefringence [39]. The reason was large scale mass transport on the azo film. A measurement demonstrating this mass transport is presented in Fig. 24. The polymer moves and no new chemical bonds are formed.

The driving mechanism behind the mass transport is not presently understood, but it is polarization-dependent [34]. The surface modulation difference between *s*- and *p*-polarization can be as much as 25 times greater for *p*-polarization than it is for *s*-polarization [40]. The height difference between the lowest and the highest point in the SRG is called the modulation depth. Circular polarization, unpolarized light and linear polarizations other than *s* and *p* have been also studied [40]. The polarization-dependence in the SRG formation was first studied by Kim *et al.* [41]. The dependence is not surprising, because Todorov *et al.* had reported polarization-dependence in the photoinduced birefringent grating formation [36].

In this work, azo-polymer SRGs were used as an etch mask. This method was first used in this manner, to the author's knowledge, by Kravchenko *et al.* [14]. The main reason why this method was chosen is that large area gratings (2 cm times 2 cm) can be inscribed with it in 90 min. This can be compared to EBL, where with normal parameters, the same area could take 56 h [42]. The method used here is an order of magnitude faster.

5.2 Nanofabrication of sub-wavelength metallic gratings

Silicon was chosen as a test material, because it is cheap and readily available in Micronova cleanroom facility. Other benefits for using silicon include the relatively easy cutting of silicon along the crystal directions that enable clear SEM pictures to be taken of the sideprofile of the gratings. Similar "easy" crystal directions do not exist in GaN,

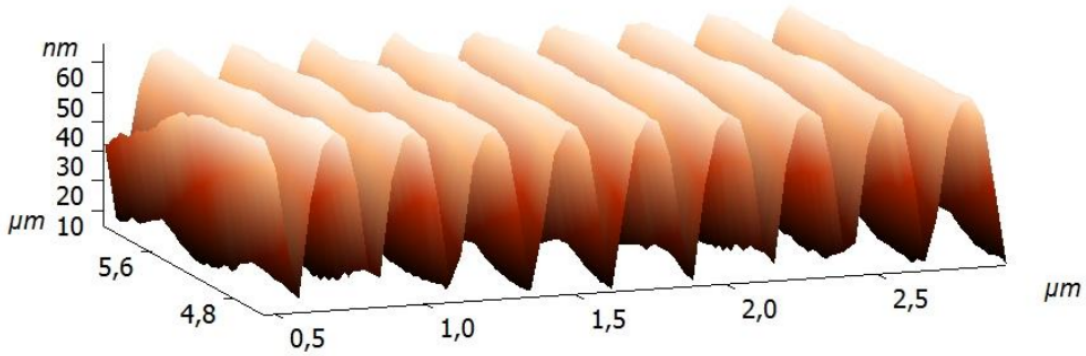

Figure 24: AFM image of a SRG formed on pDR1a. The sinusoidal shape of the polymer layer matches the values calculated from Eq. 25 .

with its wurtzite crystal structure.

5.2.1 Fabrication process

The process is very similar for Si and GaN, and also for all other smooth substrates. The first step is making crystal samples of desired size. The GaN-on-sapphire wafer was cut into quarters from a 2 inch wafer, and the silicon wafers were cut into 2 cm x 2 cm pieces. The samples were washed with deionized water after cutting to prevent the crystal dust from ruining the process.

The cut samples were coated with 40 nm of silver using electron beam evaporation (in vacuum). Silver oxidizes extremely easily. To prevent this, a 20 nm layer of gold was evaporated on the silver layer for protection. The error in the evaporation thickness is systematic, mostly due to the error in the placement of the sensor inside the vacuum chamber, and its magnitude is a few nanometers [43]. This level of error in thickness is tolerable.

Fig. 25 shows the fabrication process of the grating. For the azo-polymer spin coating, a 2 wt% solution of pDR1a in 1,2-dichloroethane is mixed, like in Ref. 14. The solvent 1,2-dichloroethane is toxic and flammable, so precautions must be taken when handling it. Then the solution is mixed in a carousel for two hours, after which it is filtered with a 0,2 μm purge filter. The solution forms clusters very quickly when exposed to air, therefore filtering and swift handling is needed. The polymer is available from Sigma-Aldrich Company Ltd. The finished solution is then spin coated on the metallized samples. The samples are spinned for 30 seconds at 5000 rounds per minute (rpm). The end result is a smooth azo-polymer film with a thickness between 90 nm and 110 nm [15]. A typical modulation depth for the exposed 1D pDR1a grating in this work is 60 nm.

When the sample has been successfully coated with metals and then spin coated, the sample is exposed to a laser interference pattern. Fig. 26 depicts the optical setup used

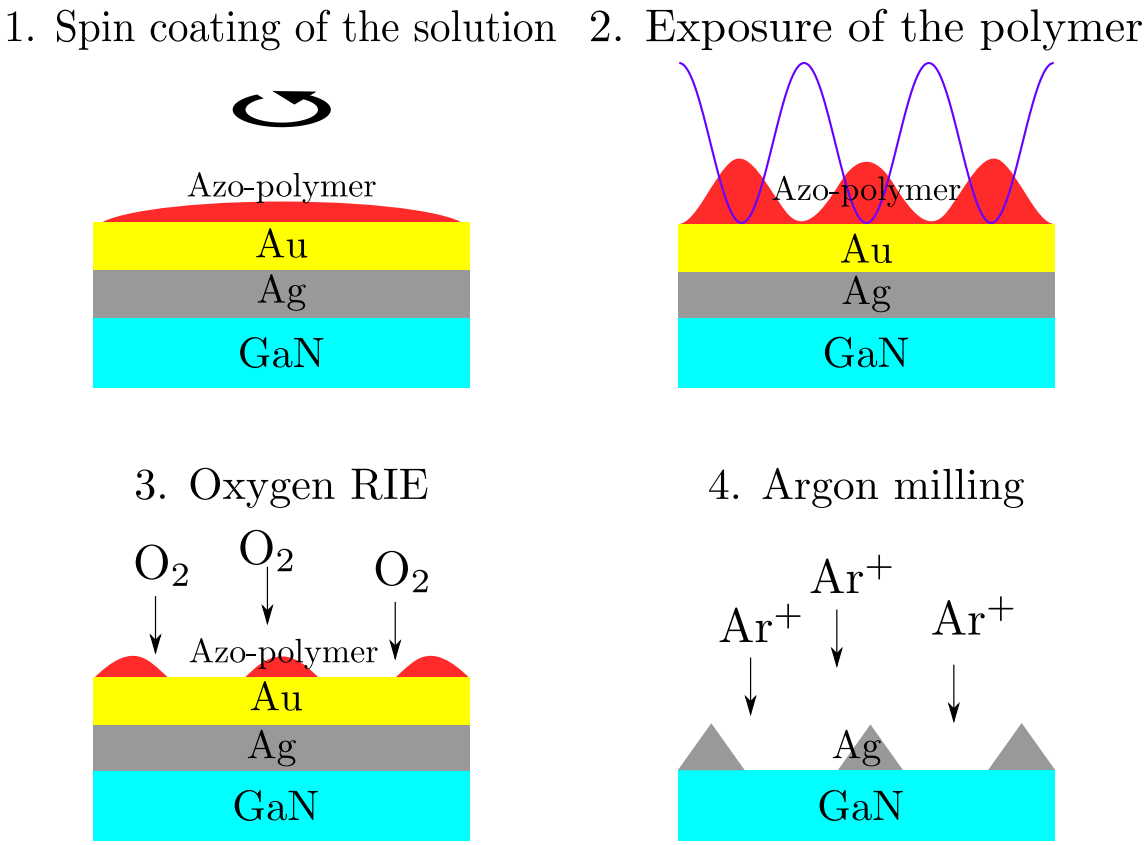

Figure 25: Main steps of the fabrication process. Step 1: the spin coating of the polymer solution. Step 2: laser interference lithography, the blue line represents the irradiance profile. Step 3: oxygen etching with RIE. Step 4: argon milling.

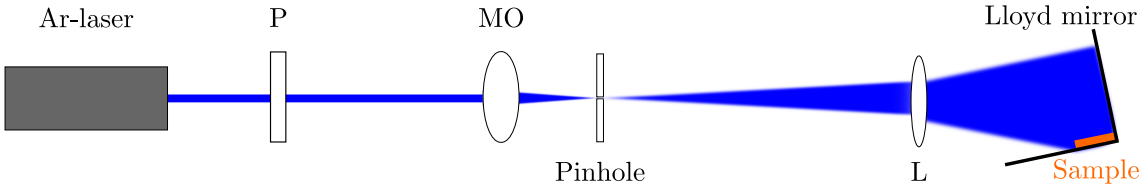

Figure 26: Schematic of the exposure set-up. The orange rectangle represents the sample to be exposed. The reflective part of the Lloyd mirror is naturally the one that the sample is not on. P is the *p*-polarizer, MO is the microscope objective, and L is a lens.

for the laser interference exposure. The microscope objective and the pinhole are known as a spatial filter (see for example Refs. 44, 45). The spatial filter cleans the laser beam profile to a near Gaussian profile. The Gaussian beam profile is smooth, but not very flat, see Fig. 28. The beam profile needs to be flattened to gain an uniform irradiance on the sample. The beam profile can be expanded by diverging the beam. This is done by placing the lens closer to the pinhole than its focal length. The focal length of the used lens is 300 mm and the distance is 275 mm. The wavelength of the laser is 488 nm. The wavelength is chosen to be near the absorbance maximum of the azobenzene group inducing repetitive *trans-cis-trans* photoisomerizations [34, 37]. The 488 nm wavelength causes significant mass transport in pDR1a as can be seen from Fig. 24. *p*-polarization, which was also used in Ref. 15, was chosen for the exposure. The angle between the incoming beam and the sample normal was 77°, see Fig. 26 and Fig. 27 . The inscribed grating period is given by

$$\Delta x = \frac{\lambda}{2 \sin \theta} \tag{25}$$

where Δx is the distance between the maxima of the irradiance (the period), λ is the wavelength and θ is the angle between the beam and the sample normal. The period can be easily changed by turning the Lloyd mirror. The Lloyd mirror is depicted in Fig. 27. The irradiance is approximately 10.5 mW/cm^2 just before the sample, and the exposure time is 90 min.

The spin coated polymer was discovered to be resistant to the final etch step by argon milling. Therefore, the azo-polymer mask is thinned with reactive ion etching (RIE). In RIE reactive gases excited by radio frequency fields are used to bombard the wafer surface [24]. Both ions and excited molecules are produced, and therefore, RIE is a combination of ionic bombardment and chemical etching. The used RIE device was a Plasmalab 80 Plus reactive ion etcher. Because the polymer is organic oxygen etching was chosen, since it produces gaseous etch products. The gold layer protects the silver against oxidation.

The oxygen etching time was used as the primary optimization parameter. Other parameters could have been chosen, for example argon milling time, solution concentration and exposure time. A longer oxygen etching etches more polymer and more gold is exposed. The width of the exposed gold stripes seems to depend linearly on the etch time, see Fig. 30. In this manner the amount of exposed metal could be controlled. 30 s cycles were

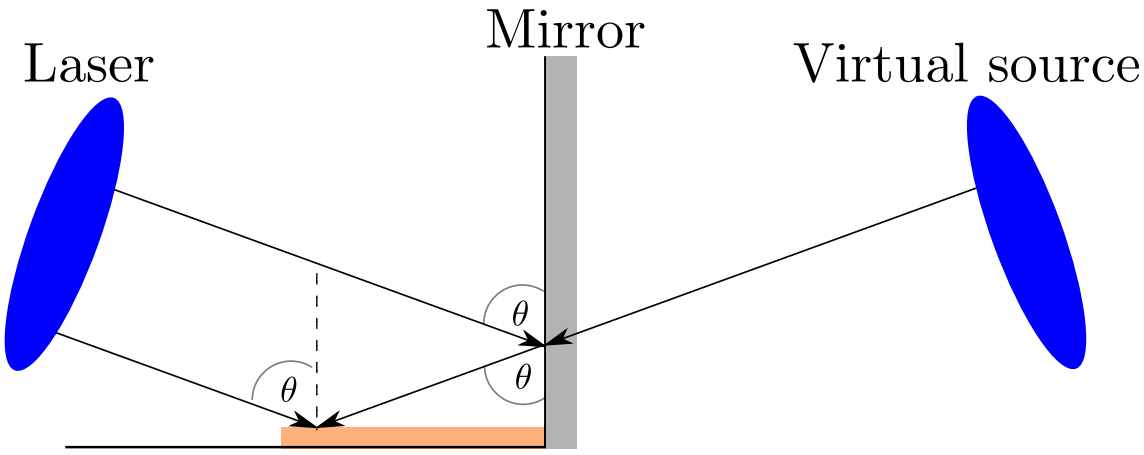

Figure 27: Illustration of the Lloyd mirror (also called Lloyd interferometer). Constructive interference occurs, when the reflected beam travels an optical path length that is an integer times the wavelength. Thus, the Lloyd mirror forms the interference pattern.

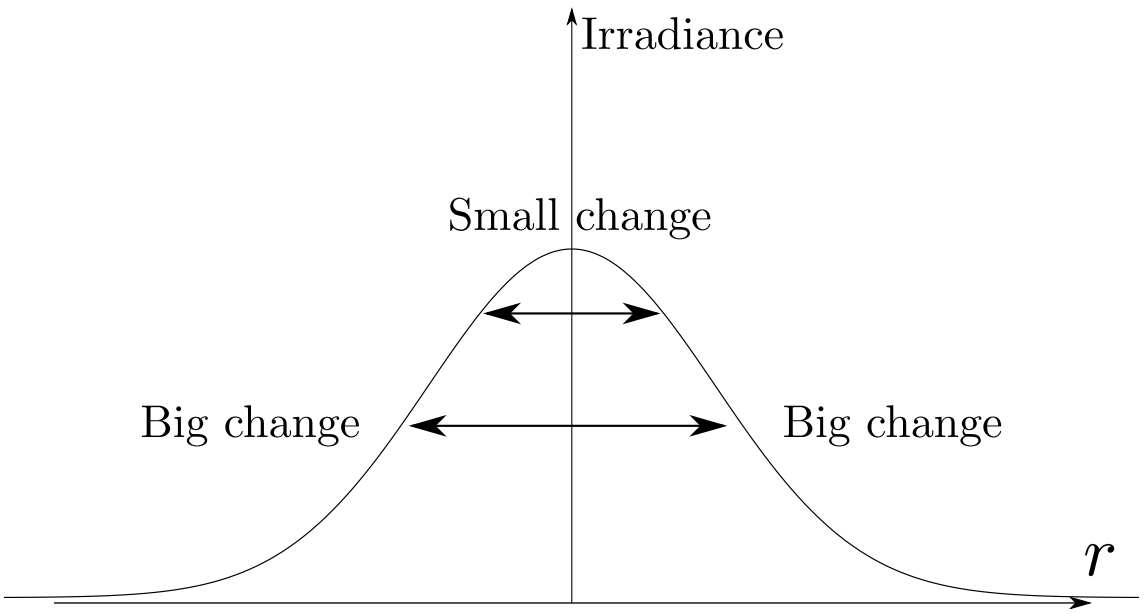

Figure 28: Gaussian beam profile of the cleaned laser beam. The change in the relative irradiance along the beam is small if only the center is used in the exposure. The change is big, if a larger portion of the beam is used. r is the distance from the beam center.

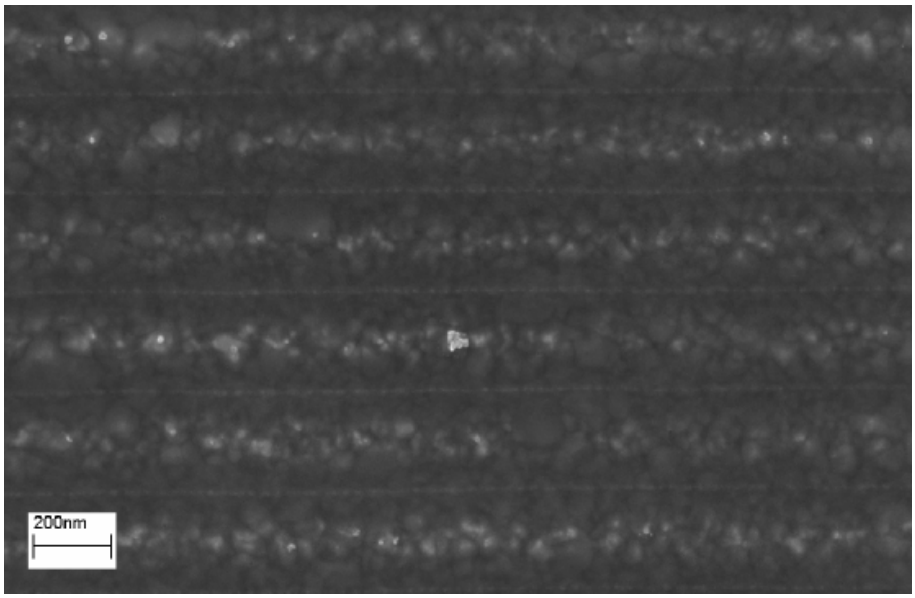

Figure 29: SEM image of azo-polymer on GaN oxygen etched for 1 min 40 s. The gold can be seen as the white glow in the valleys. The thin white stripes are the top of the polymer. The white colour in the polymer tops is due to the edge effect in SEM.

used initially, and if more etching was needed, a cycle of 40 s was used. The pressure during the etch is 1.95 Pa, the radio frequency power is 40 W and the oxygen flow is 40 standard cubic centimeters per minute. A SEM image of the etched polymer sample can be seen in Fig. 29, showing the ideal mask for grating fabrication. The gold is just beginning to show as white in the image. In SEM the backscattered electrons are detected from the primary electron beam, and this backscatter is strongest when the beam hits metals or thin high points (the edge effect). So when gold is barely covered by the polymer the minimum oxygen etch time is found. By etching more of the polymer different widths of the grating stripe can be attained.

The last step in the making of the gratings is argon milling. Argon atoms are ionized in the RIE chamber and the ions chip away any material on their way. The ion milling was done in the same Plasma Lab 80 Plus reactive ion etcher as the oxygen ething. Cycles of 30 s exposures were used again. The pressure during the etch is 1.95 Pa, the radio frequency power is 100 W and the argon flow is 25 standard cubic centimeters per minute. A photograph of one of the finished GaN samples is seen in Fig. 31. SEM images of finished gratings are displayed in Figs. 32 to 35.

A surprisingly triangular profile was generated in the manufacturing process, see Figs. 32 and 36. The optimal recipe was found to consist of 1 min 40 s oxygen etching and 16 cycles of Ar milling. This produces gratings whose stripe width is 130 nm. The period is approximately 250 nm in all the samples. The aim was to make square profiled gratings with a width of 150 nm and a period of 250 nm like in Ref. 11. Due to the triangular profile, this was not achieved. It was deemed sufficient to make gratings with the same

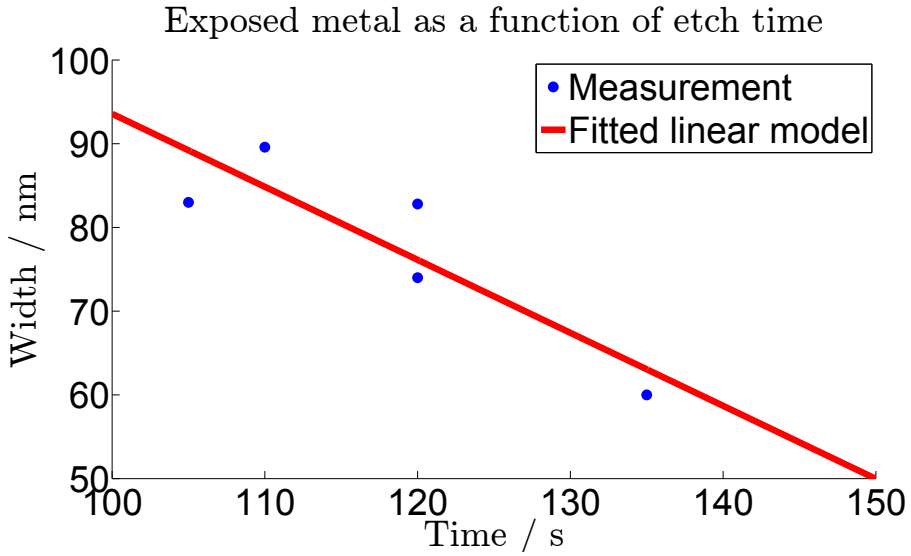

Figure 30: Width of the exposed metal stripes as a function of oxygen etch time. The width of the exposed gold stripes seems to be a linear function of the etch time. Note that this graph represents the situation without Ar milling.

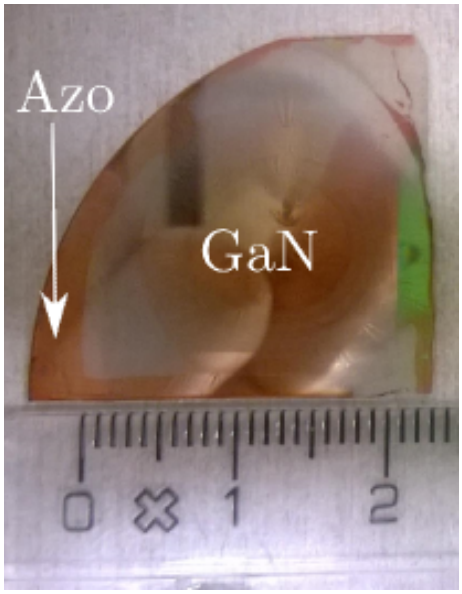

Figure 31: Photograph of the same sample as in Fig. 35. A ruler for scale. The author's finger and green camera are reflected from the GaN surface. The red colour comes from the azo-polymer.

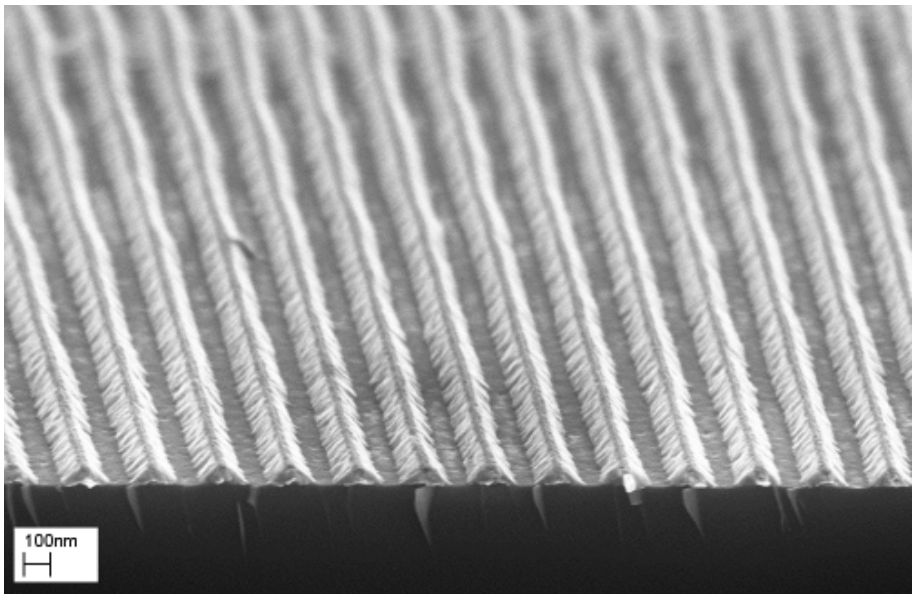

Figure 32: Cross section SEM image of a silver grating with 170 nm stripe width on silicon. Made with 1 min 50 s oxygen plasma etch and 18 cycles of argon milling. Residual azo-polymer is visible on the top of each stripe. This test sample was made with an old polymer solution, compare to Fig. 38.

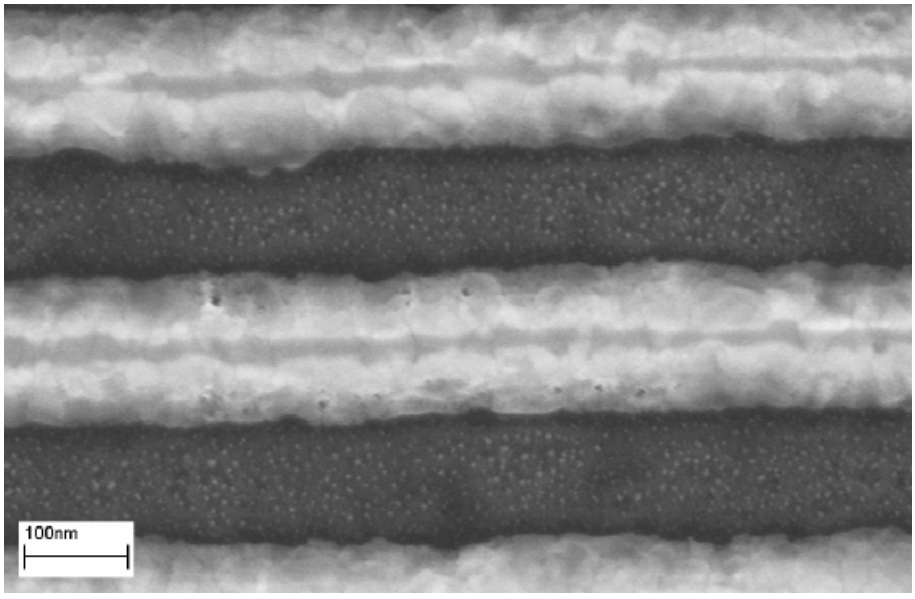

Figure 33: Close up from the top direction of the same grating as in Fig. 32.

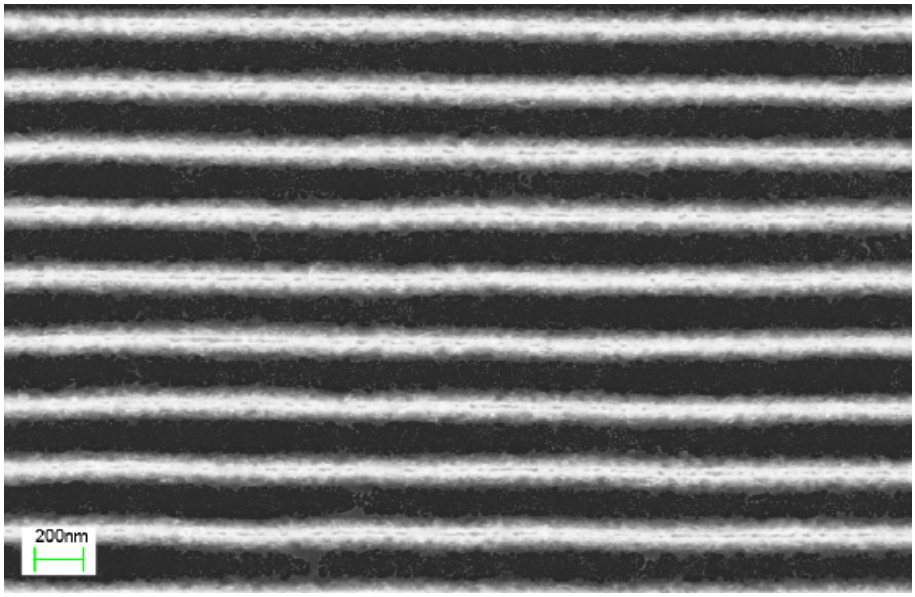

Figure 34: Silver grating on silicon. Oxygen etch time was 1 min 40 s and 14 argon milling cycles were used.

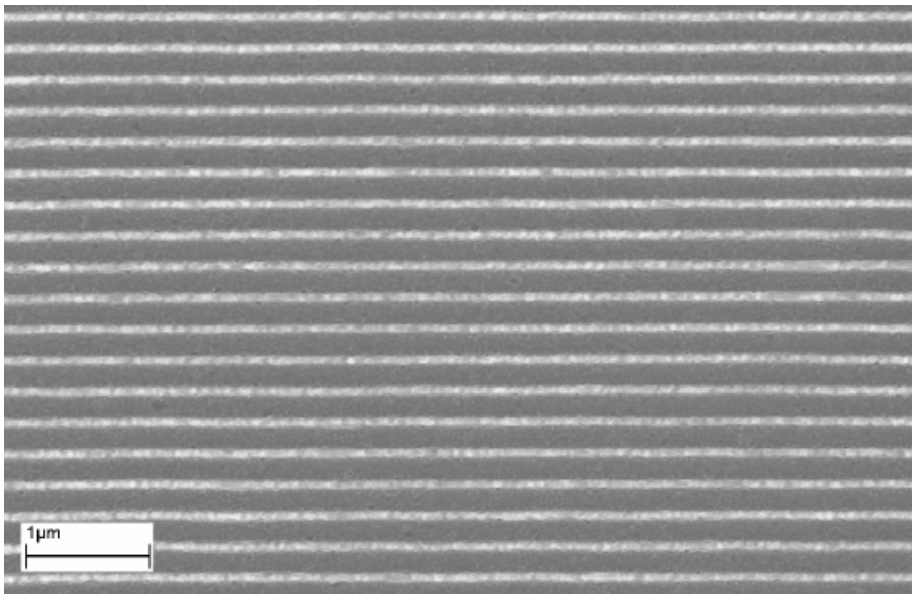

Figure 35: Silver grating on GaN. Oxygen etch time was 1 min 40 s and 16 argon milling cycles were used. The silver stripe width is 130 nm.

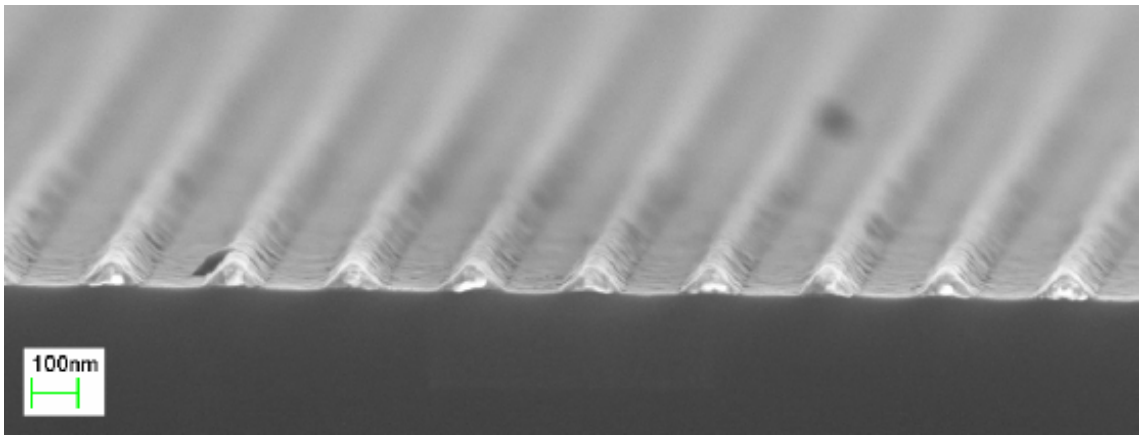

Figure 36: SEM image of a silver grating on silicon. Oxygen etch time was 1 min 40 s and 16 argon milling cycles were used. The stripe width is 130 nm.

period, but with the thinner 130 nm width. The manufacturing process was very similar to Refs. 15 and 14. In both articles, the final structures had rectangular side profiles with a slightly curved upper part. The reason for this different behaviour is unknown at the moment. The lower height of the structures and the thick gold sacrificial layer might play a role. Also a larger period might allow a relatively flatter top, because a larger modulation depth can be attained [34].

If this technique is to be used in a LED fabrication process, it must be scalable. Therefore whole wafers need to be processed in one run. RIE processes used here are already done on whole wafers. The remaining limiting factors to the sample size are the spin coater and being able to expose the whole sample with the laser. Because the spatial filter loses a large portion of the beam energy, a powerful laser or a longer exposure time is needed. Two lasers can be used to offset this power loss. In this confuguration no Lloyd mirror is used. The two laser beams form the interference pattern. Ultimately the power loss can be compensated, either with two lasers, longer exposure time or one very powerful laser. Thus, the process presented here is scalable.

5.2.2 Challenges

Several problems occurred during the process development. First, the used gold layer was not thick enough. The initial recipe had a 10 nm gold layer. The silver oxidized during the oxygen etching through the combined polymer and gold. This can be seen in Fig. 37. When the gold layer thickness was 20 nm, the problem disappeared, as can be seen from, *e.g.*, Fig. 35.

Another problem was found when processing silicon samples: the age of the azo-polymer solution has a significant impact on the mass transport. Old polymer solutions had much smaller modulation depths. This naturally affects the end result of the RIE process. The same recipe that produced the grating seen in Fig. 32 produced a completely overetched

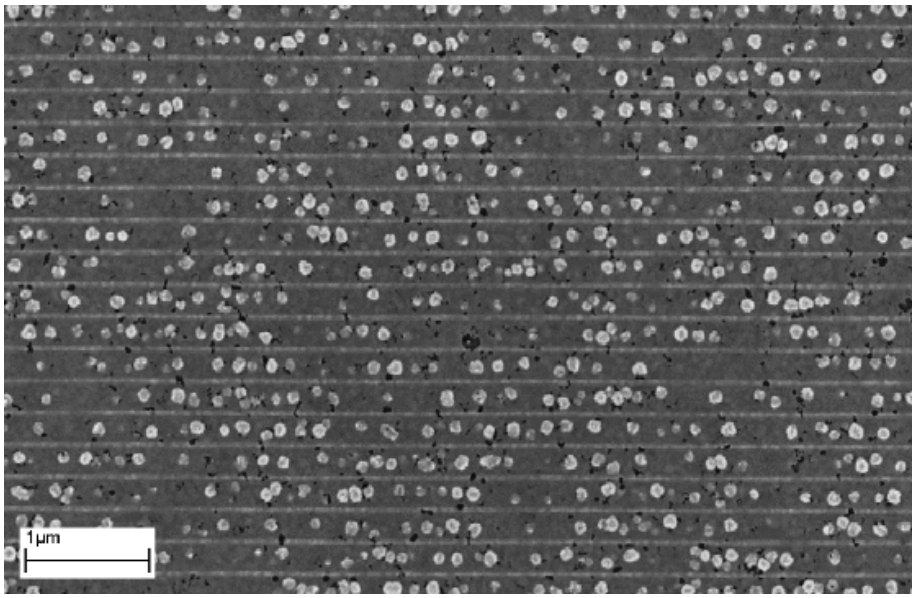

Figure 37: Granules seen in this SEM picture are oxidized silver. Silver stripes on silicon can also be seen.

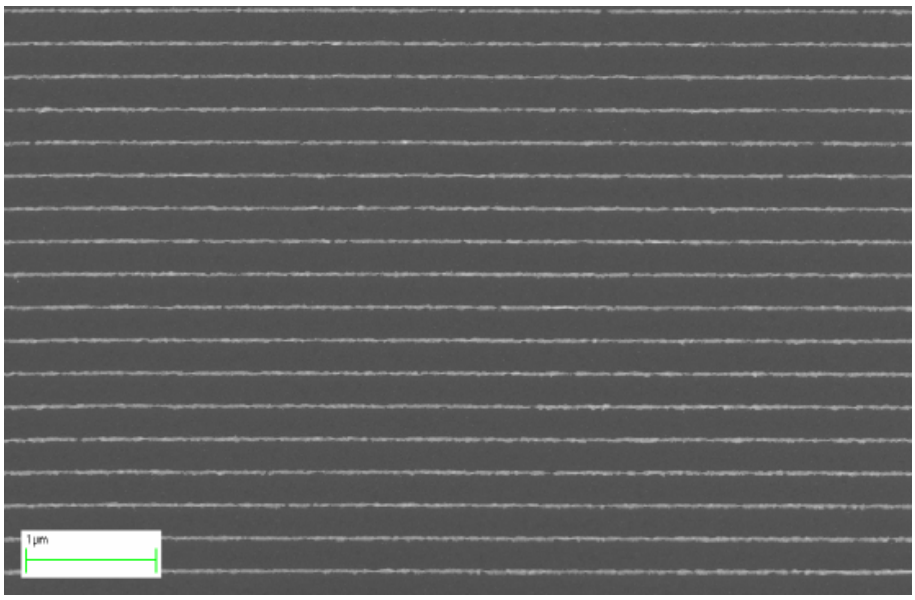

Figure 38: SEM image of a silver grating on silicon. Made with the same recipe, but two weeks younger solution than the grating in Fig. 32.

grating seen in Fig. 38. Note that this recipe was not the optimized recipe. The latter grating was made with a two week old solution. It was seen prudent after this that solutions should not be used after a week, and preferrably they should be used within a few days after mixing and filtering.

The smoothness of the gold layer is also critical in the process. The electron beam evaporator sometimes produced large gold spheres on the surface. The particles were tens of

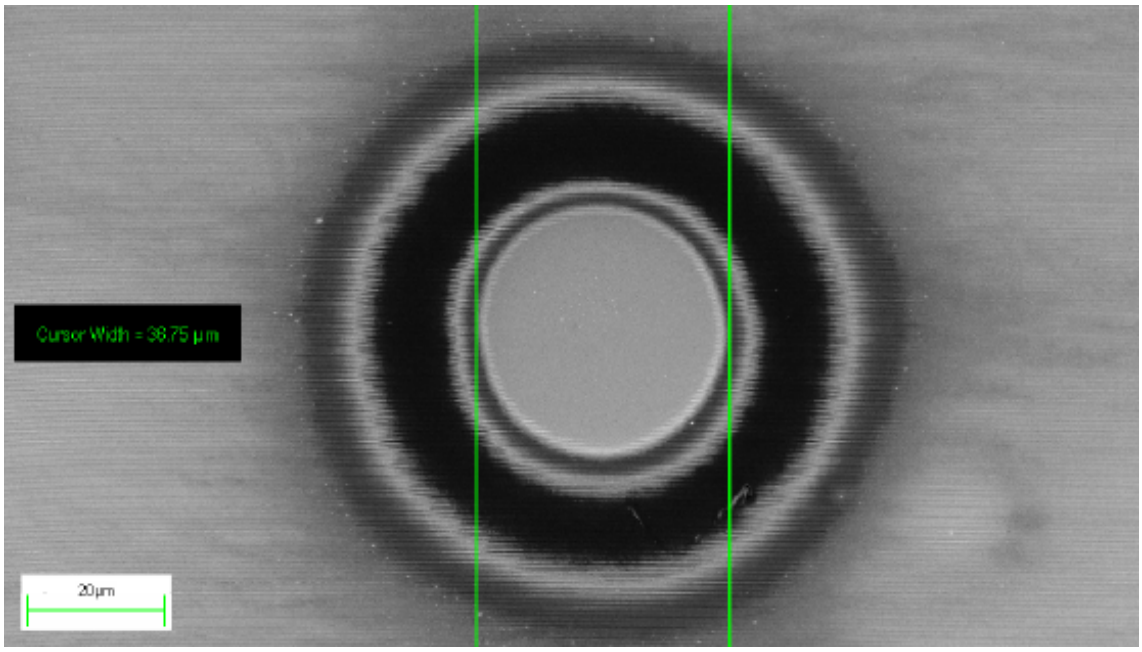

Figure 39: One of the gold spheres formed in the evaporation. Picture taken with SEM after all the normal processing.

micrometers in diameter. When the spinned azo-polymer layer is approximately 100 nm thin, these spheres significantly alter the spin coating process [15]. One such sphere can be seen in Fig. 39 . Thousands of these could be on one sample, rendering the properties of the grating unreliable. During the later stages of this work their formation stopped for unknown reasons.

This technique encompasses several positive qualities. It allows large areas to be patterned quickly. Using EBL to make these features on a macrosopic scale (the same 2 cm x 2 cm as the samples here) would be extremely expensive, slow and detrimental to the QW [16]. Also, the chemical properties of the substrate do not matter. No chemical etching is done to the substrate, which permits a wide variety of substrates to be used. Because the gold protects the silver, and the final etching is argon milling, the hardness of the substrate is the only issue in the milling. The harder the substrate, the less it is damaged. No significant damage was observed during this work, though. The optical properties of the substrate are not relevant either, because during the exposure the substrate is covered with a 60 nm thick noble metal layer. Visible light is highly evanescent in such media. It would not be surprising if this technique would find niche uses in the semiconductor industry.

5.2.3 2D gratings

2D azo-polymer gratings can be inscribed with the same exposure setup previously presented. This is done by first inscribing a 1D azo-polymer grating on the sample at 9.0 mW/cm^2 for 90 min. The sample is then rotated 90° and a shorter 13 min exposure

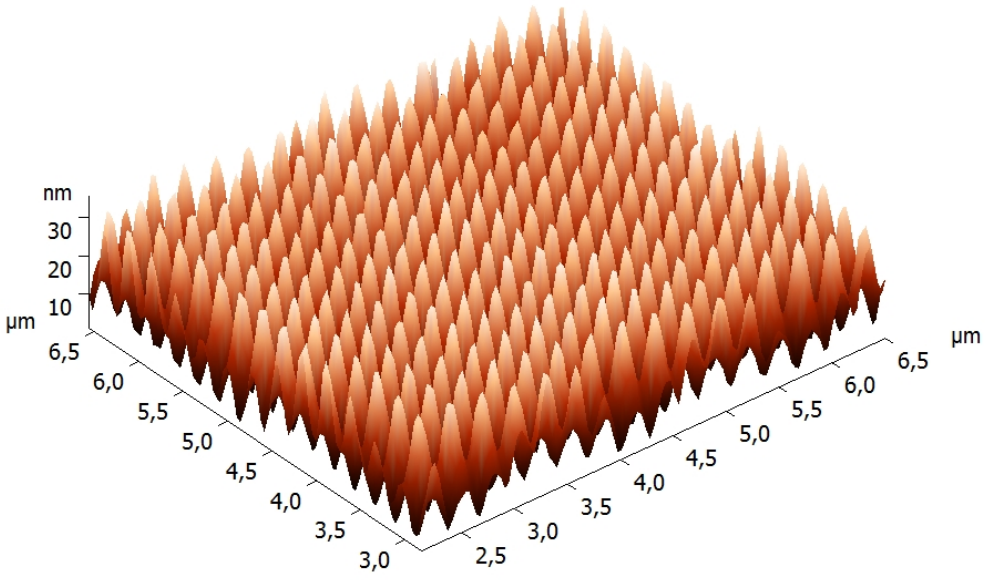

Figure 40: AFM image of spin coated azo-polymer exposed two times with the second interference pattern being orthogonal to the first. A 2D-grating emerges. 3D data is tilted for viewing purposes.

is made with the same irradiance. 2D gratings can be made this way from the polymer, see Figs. 40 and 41. Fig. 42 demostrates the repeatability of the 2D azo-polymer process. Image of the polymer in the middle oxygen etching is seen in Fig. 43. These 2D SRGs can also be used as etch masks.

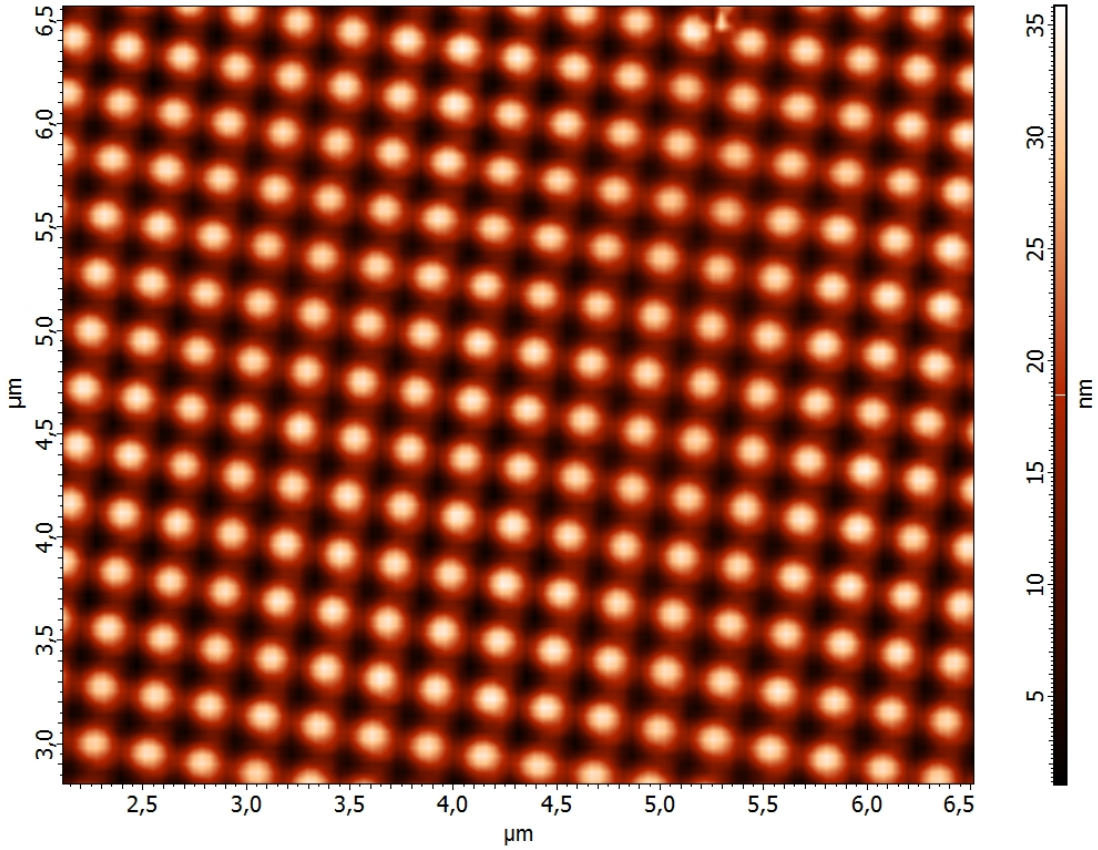

Figure 41: Top view of the same AFM scan as in Fig. 40.

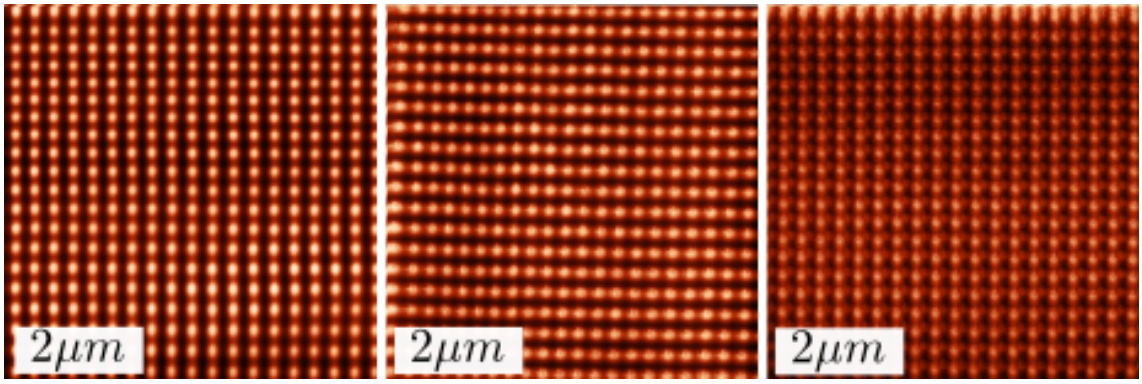

Figure 42: Three separate samples with the same polymer concentration and the same exposure times were measured using AFM. Sample on the left and in the middle have a modulation depth of 35 nm and the rightmost sample has a modulation of approximately 32 nm. Because of the slight asymmetry of the exposed polymer, the modulation depth was calculated in the direction of the highest modulation.

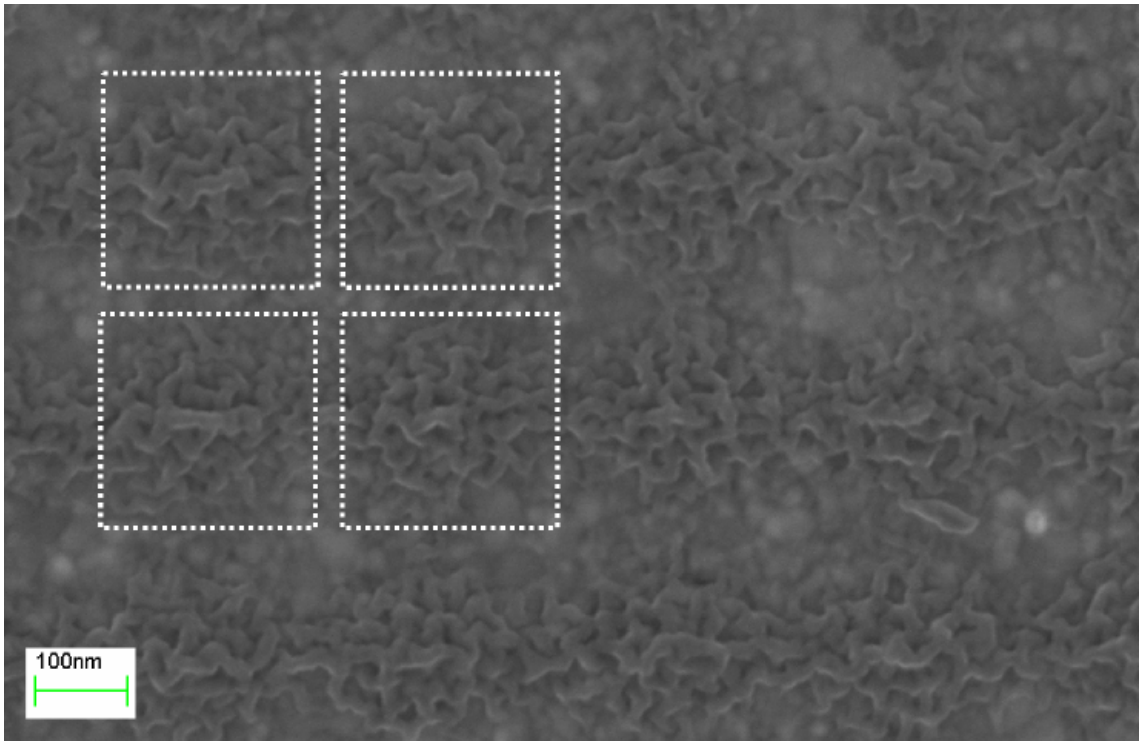

Figure 43: SEM image of oxygen etched azo-polymer on gold coated silicon. Two orthogonal exposures were performed. The 2D structure of the azo-polymer can clearly be seen. The dashed rectangles are centered each on a polymer "hill".

6 Results

This section focuses on the results. In this work the amount of light emitted through SP coupling was measured from InGaN QW samples with a silver grating. A summary of the development of silver gratings on GaN is presented. Periodic arrays of gold spheres were also fabricated. The gold sphere arrays and nanopillars grown from them are discussed. Finally the SP coupling experiment and its results are presented.

6.1 Summary of subwavelength silver gratings on GaN

Silver gratings were manufactured on GaN, see *e.g.* Fig. 35. The development process of the gratings and the finished gratings are discussed extensively in Chapter 5. The purpose of the gratings was to couple energy from an InGaN QW to SP modes on the gratings, and then couple these SPs to the radiative modes. A total of six samples with the discussed 130 nm stripe width were made. The period and height of the gratings were, respectively, 255 nm and 40 nm. The gratings showed uniformity between the samples and within the samples. Significant damage to the substrates was not observed.

The azo-polymer etch mask process is an order of magnitude faster than EBL [42]. This was seen when fabricating the gratings. Large area (2 cm x 2 cm) subwavelength gratings could be exposed in 90 min compared to the 56 h with EBL.

6.2 Periodic nanostructures for patterning and nanopillar growth

In addition to the silver gratings, arrays of gold nanostructures were made. This was done to develop the azo-polymer process further. 2D azo-polymer etch masks were used to produce the structures in Fig. 44. The gold arrays were made using an azo-polymer etch mask to shape a gold thin film into a grid of gold islands. The sample seen in Fig. 45 is made from a structure like that shown in Fig. 44 by heating the sample. The gold islands melt and form spheres, which then crystallize when cooled. The size of the gold particles is determined by the amount of milling cycles used, and the period is determined by Eq. 25. This means that the gold sphere diameter is completely controlled by varying the amount of milling cycles. The uniformity of the spheres is rather good based on several SEM images of the arrays.

The purpose for these arrays of gold spheres was to grow GaAs nanopillars in MOVPE using these gold spheres as catalysts. A very similar process has been demonstrated in Ref. 46. The details about the nanopillar growth in MOVPE are beyond this work. The end result from the MOVPE growth is presented in Fig. 46. The grown nanopillars are ordered in the sites of the gold spheres. Nanopillars have many interesting properties [17, 19, 46], and they can be used to make nanopillar solar cells [46].

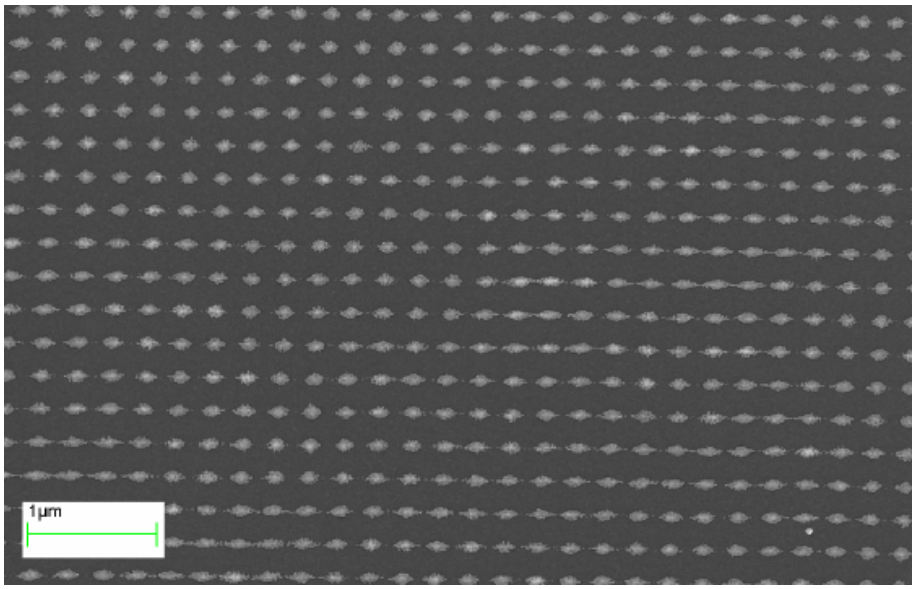

Figure 44: SEM image of a periodic array of gold islands on silicon, manufacured with an azo-polymer mask.

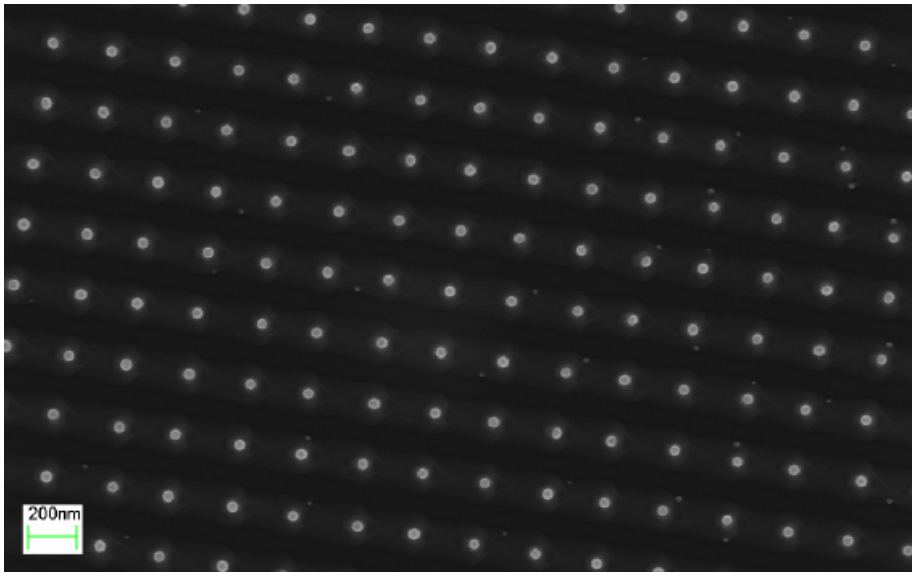

Figure 45: SEM image of a periodic array of gold spheres on GaAs, manufacured with an azo-polymer mask. The sphere diameter is 40 nm. Image taken by Tuomas Haggren.

6.3 Surface plasmon coupling experiment

Polarization ratio (PR) is used to estimate the amount of coupled light. The coupled light is due to SP modes coupling to radiative modes. The PR is defined as the power in the vertical polarization divided by the power in the horizontal polarization, see Fig. 18. Unpolarized light would have a PR of one. The PR is a good quantity for this purpose because the light from the SP is on average vertically polarized (see Chapter 4). If a significant part of the radiation is due to SPs from strong quantum electrodynamic coupling,

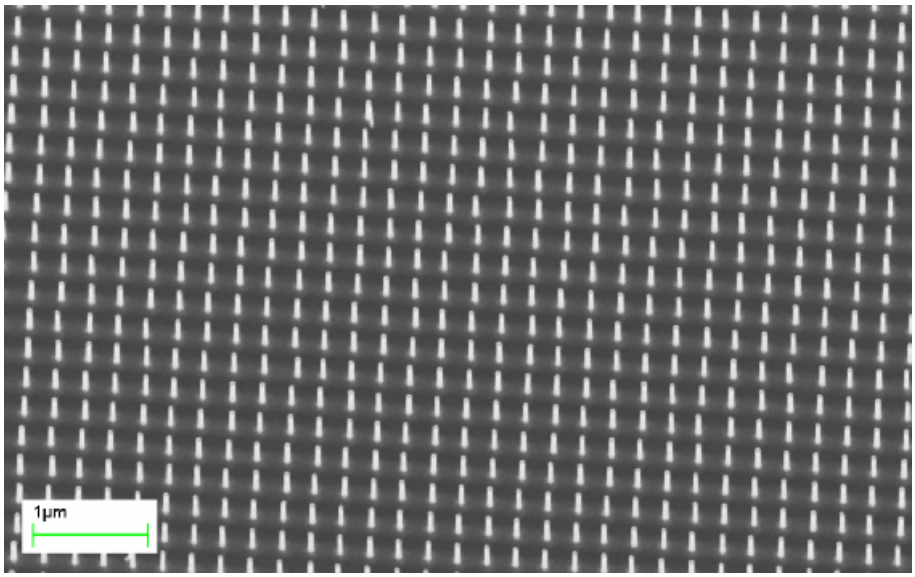

Figure 46: Periodic array of GaAs nanopillars on GaAs, grown from the sample in Fig. 45. Image taken by Tuomas Haggren.

then the PR should drop exponentially when increasing the capping layer width. This is due to the exponential dependence of the strong quantum electrodynamic coupling on the capping layer width (see Chapter 4). In short, the PR can be used as a measure of the amount of light radiating from SP modes.

The classical coupling is not sensitive to the capping layer width, therefore the PR should not drop to 1 as the width is increased, because the classical coupling will produce on average vertically polarized radiation. The amount of classically coupled SP ligth and non-classically coupled SP ligth can be separated. When the QW is positioned deep enough in the GaN to render the strong quantum mechanical coupling neglible, the remainder of the PR is from the classical SP coupling. If the PR drops to 1 it means that the classical coupling is weak.

The PR experiment itself is simple, even if rather experimentally tedious (see Fig. 47). The QW is pumped with a $\lambda = 325\,\text{nm}$ He-Cd UV laser from the epitaxy side. The laser is a Plasma model HCGL-15UM. A plano-convex lens is used to focus the laser on to the sample. Front side excitation is used because the processed wafers are single side polished, and backside excitation through the rough milky white sapphire is nearly impossible. Glancing angle excitation is used to minimize any radiation from the laser to the spectrometer. This is done because the laser spectrum contains several peaks in the visible region. In a glancing angle excitation, the exciting laser beam is reflected from the sample surface. The photoluminescence (PL) from the sample is gathered with a high numerical aperture (NA) microscope objective, while the reflected beam passes the microscope objective (Fig. 47). This way no optical filters are required to separate the excitation from emission, improving the signal-to-noise ratio. The gathered PL is focused to an optical fiber with a plano-convex lens. The fiber is connected to an Ocean optics USB 2000

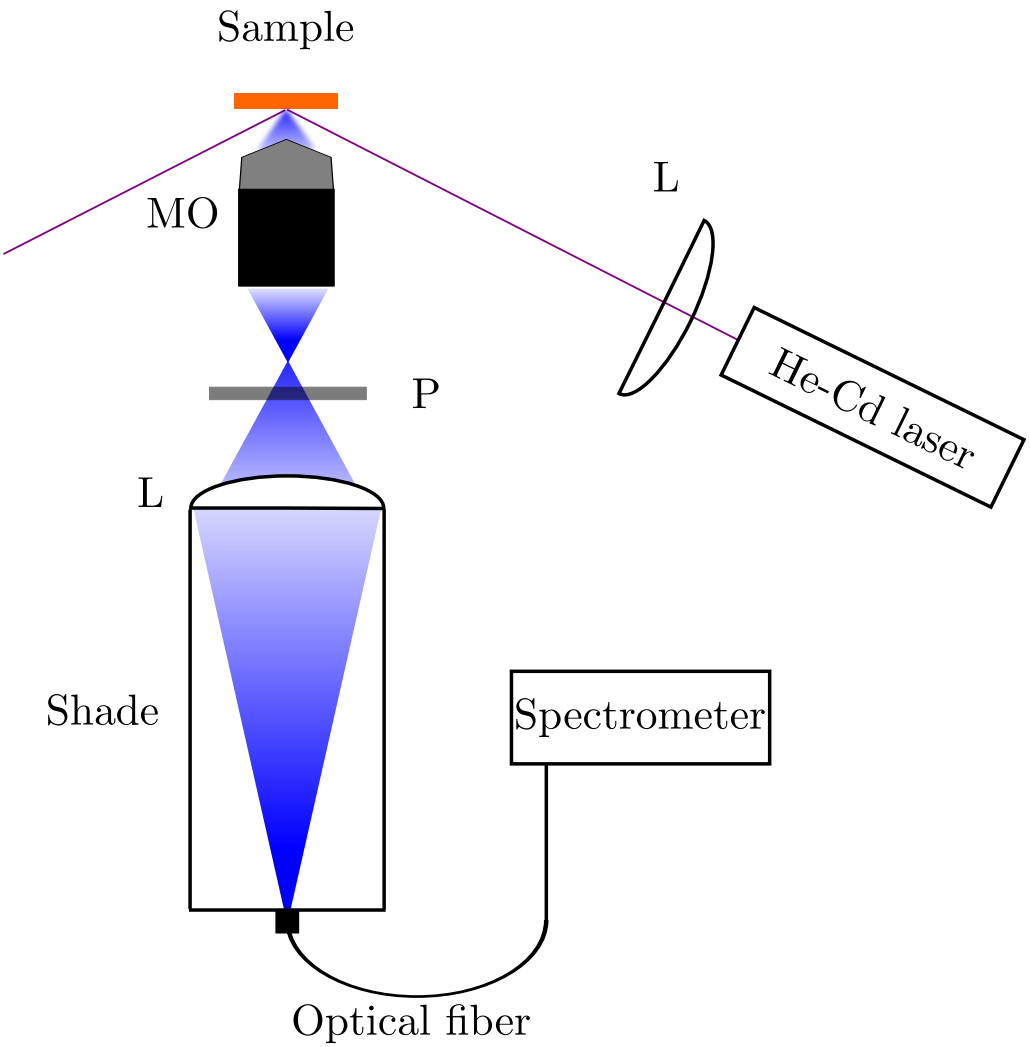

Figure 47: Plasmon coupling experiment. The QW is excited with a glancing angle focused laser beam. L is a lens, MO is a microscope objective and P is a rotatable linear polarizer.

spectrometer. The two polarizations are measured by putting a rotatable linear polarizer (Thorlabs LPVISE dichroic film polarizer) between the microscope objective and the second lens. In the experiment a shade tube is used to minimize any stray light hitting the fiber end from the surroundings or from the scattered laser light from the sample surface (Fig. 47).

A model based on the strength of the non-classical coupling is compared with the measurements. The model describes the strong quantum electrodynamic coupling without the classical SP coupling. The model is of the form

$$PR = ae^{-2k_D \cdot w} + 1 \tag{26}$$

where the term $e^{-2k_D \cdot z}$ comes from Eq. 24, w is the capping layer width, k_D is the GaN wavevector at 470 nm. The "+1" in Eq. 26 comes from the fact that the PR approaches

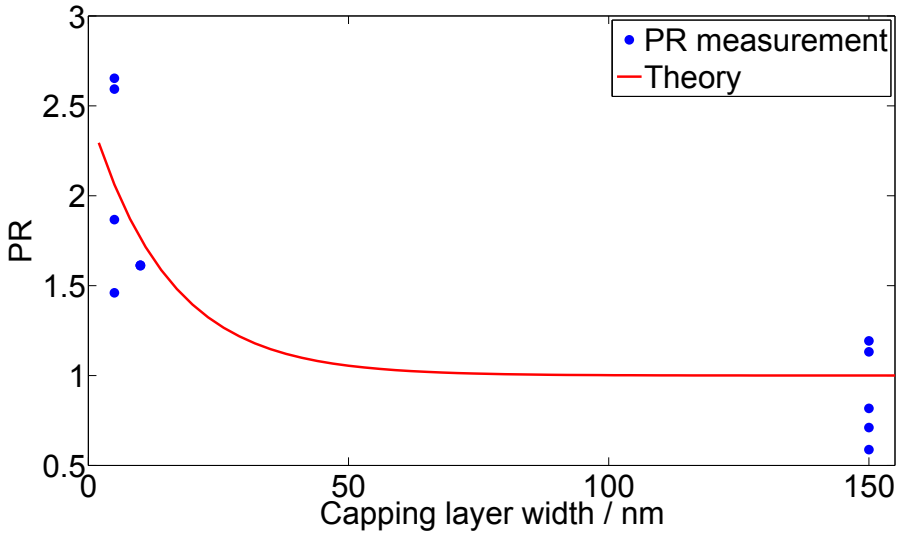

Figure 48: Results from the SP coupling experiment. The theoretical curve estimates the amount of strong quantum electrodynamic coupling. 12 PR measurements are shown (some are very close to each other).

one at large w. The coefficient a estimates the dipole moment of the electron-hole pair squared times the plasmon density of states at the band gap energy, see Eq. (24). Calculating a from theory is challenging, but because it is just a scaling factor it can be estimated by fitting the model to measurement data.

A qualitative analysis of the data in Fig. 48 shows that the theory fits the measurement data reasonably well. The PR seems to drop exponentially. This suggests that the QW energy might be coupled to SP modes, which then radiate. More data points are needed between capping layer thicknesses 10 nm and 50 nm to support that the PR is actually due to strong quantum electrodynamic coupling. Presently the data remains inconclusive with respect to the validity of the model.

It is odd that the PR drops to 1 as the capping layer width is increased, because the classical SP coupling should not disappear. If the three lowest data points at 150 nm are discarded, the mean PR at that point is 1,16. Anything over one is interpreted as the classically coupled SP radiating. This number would suggest that the classical coupling is weak even if the three points are discarded. However, because the three points are in a cluster, they can not be dismissed. The cluster might still be due to measurement error. The two other points are in line with expectations at 150 nm.

An alternative theory to SPs as the cause for the PR is that the grating acts as a wire grid polarizer. However, because the PR drops when the QW is deeper in the sample, this does not support the wire grid polarizer hypothesis. If the silver grating would act as a polarizer, then the PR should not depend on the distance between the QW and the grating. Therefore, the grating probably does not act as a polarizer.

Similar measurements of SP coupling are presented by Shen *et al.* [10]. They concluded that the PR of the radiation was due to strong quantum electrodynamic coupling. These experiments actually provided the inspiration for using PR as the measure of SP coupling in this work. Although the PR of the electroluminescence was studied in Ref. 10, the results are comparable. A PR of approximately 1,6 was common in Ref. 10, which does not significantly differ from the values that were measured for this work (Fig. 48). Similarly to this work the PR was measured from an InGaN QW that had a silver grating. The grating period was 500 nm and the height was 30 nm [10], which differs from the values used in this work.

7 Conclusions

In this work large area (2 cm x 2 cm) subwavelength silver gratings were fabricated on GaN. The aim with these gratings was to couple energy from an InGaN QW inside the GaN to SP modes and subsequently to couple these to radiative modes. The gratings were fabricated using azo-polymer SRG etch masks, which were made using laser interference lithography and RIE.

It was shown that azo-polymer SRG can efficiently be used to fabricate a variety of nanostructures on different substrates. Subwavelength silver gratings on GaN and silicon, and arrays of gold spheres on silicon and GaAs were created using azo-polymers. Large area fabrication of all of these nanostructures was demonstrated. The process was shown to be repeatable, but very sensitive to the age of the mixed azo-polymer solution. The developed silver grating fabrication process showed uniformity within samples and between samples. No significant damage was observed on the GaN or silicon substrates. Likewise, no damage to the InGaN QW was observed unlike in EBL fabricated gratings. The fabricated arrays of gold spheres on GaAs were used in MOVPE to produce ordered arrays of GaAs nanopillars.

SP coupling experiments were performed with InGaN QW inside GaN with the silver grating being on top of the sample. The built experiment was simple and relatively easy to perform. Moreover, the experiment is not sensitive to the laser containing wavelengths in the measured wavelength region. The theory fits the measurement data qualitatively. The emission from the QW might be strongly coupled to SP modes, but additional experiments are needed to verify this.

To conclude, azo-polymer masks seem promising for several applications in nanotechnology. It is possible that azo-polymer masks find specific uses in the semiconductor industry, *e.g.*, for nanopillars and subwavelength gratings. For the SP coupling, more study is needed, but at least for nanofabrication purposes azo-polymer masks are suitable. This is demonstrated by the numerous structures and their SEM images in this work. The time seems ripe for azopatterning.

References

[1] W. Paul and T. Satoshi, *Light's Labours Lost: Policies for Energy-efficient Lighting* (International Energy Agency, 2006).

[2] S. Nakamura and G. Fasol, *The Blue Laser Diode: GaN Based Light Emitters and Lasers* (Springer-Verlag, 1997).

[3] M. Craford, "From holonyak to today," Proceedings of the IEEE **101**, 2170–2175 (2013).

[4] C. Huh, K. Lee, E. Kang, and S. Park, "Improved light-output and electrical performance of InGaN-based light-emitting diode by microroughening of the p-GaN surface," Journal of Applied Physics **93**, 9383–9385 (2003).

[5] S. Chang, L. Chang, D. Kuo, T. Ko, S. Hon, and S. Li, "GaN-based LEDs with rough surface and selective KOH etching," Display Technology, Journal of **10**, 27–32 (2014).

[6] M. Kwon, J. Kim, I. Park, K. S. Kim, G. Jung, S. Park, J. Kim, and Y. Kim, "Enhanced emission efficiency of GaN/InGaN multiple quantum well light-emitting diode with an embedded photonic crystal," Applied Physics Letters **92**, 251110 (2008).

[7] S. Patra, S. Adhikari, and S. Pal, "Design and analysis of "chess board" like photonic crystal structure for improved light extraction in GaN/InGaN LEDs," J. Display Technol. **9**, 339–345 (2013).

[8] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, 2012), 2nd ed.

[9] D. Yeh, C. Huang, C. Chen, Y. Lu, and C. Yang, "Surface plasmon coupling effect in an InGaN/GaN single-quantum-well light-emitting diode," Applied Physics Letters **91**, 171103 (2007).

[10] K. Shen, C. Chen, H. Chen, C. Huang, Y. Kiang, C. Yang, and Y. Yang, "Enhanced and partially polarized output of a light-emitting diode with its InGaN/GaN quantum well coupled with surface plasmons on a metal grating," Applied Physics Letters **93**, 231111 (2008).

[11] E. Homeyer, P. Mattila, J. Oksanen, T. Sadi, H. Nykänen, S. Suihkonen, C. Symonds, J. Tulkki, F. Tuomisto, M. Sopanen, and J. Bellessa, "Enhanced light extraction from InGaN/GaN quantum well with silver gratings," Applied Physics Letters **102**, 4 (2013).

[12] D. Yeh, C. Huang, C. Chen, Y. Lu, and C. Yang, "Localized surface plasmon-induced emission enhancement of a green light-emitting diode," Nanotechnology **19**, 345201 (2008).

[13] I. Gontijo, M. Boroditsky, E. Yablonovitch, S. Keller, U. K. Mishra, and S. P. DenBaars, "Coupling of ingan quantum-well photoluminescence to silver surface plasmons," Phys. Rev. B **60**, 11564–11567 (1999).

[14] A. Kravchenko, A. Shevchenko, V. Ovchinnikov, A. Priimagi, and M. Kaivola, "Optical interference lithography using azobenzene-functionalized polymers for micro- and nanopatterning of silicon," Advanced Materials **23**, 4174–4177 (2011).

[15] A. Kravchenko, A. Shevchenko, P. Grahn, V. Ovchinnikov, and M. Kaivola, "Photolithographic periodic patterning of gold using azobenzene-functionalized polymers," Thin Solid Films **540**, 162–167 (2013).

[16] H. Nykänen, P. Mattila, S. Suihkonen, J. Riikonen, E. Quillet, E. Homeyer, J. Bellessa, and M. Sopanen, "Low energy electron beam induced damage on InGan/Gan quantum well structure," Journal of Applied Physics **109** (2011).

[17] H. Ibach and H. Lüth, *Solid-state physics* (Springer, 2009), 4th ed.

[18] M. Born and R. Oppenheimer, "Zur Quanthentheorie der Molekeln," Annalen der Physik **84**, 457–484 (1927).

[19] P. Bhattacharya, *Semiconductor optoelectronic devices* (Prentice-Hall of India, 2005), 2nd ed.

[20] S. Nakamura, T. Mukai, and M. Senoh, "Candela-class high-brightness InGaN/AlGaN double-heterostructure blue-light-emitting diodes," Applied Physics Letters **64** (1994).

[21] J. Yang, C. Sun, Q. Chen, M. Anwar, M. Asif Khan, S. Nikishin, G. Seryogin, A. Osinsky, L. Chernyak, H. Temkin, C. Hu, and S. Mahajan, "High quality GaN-InGaN heterostructures grown on (111) silicon substrates," Applied Physics Letters **69**, 3566–3568 (1996).

[22] O. Madelung, ed., *Semiconductors: group IV elements and III-V compounds* (Springer-Verlag, 1991).

[23] M. Bass, ed., *Handbook of Optics Volume IV* (McGraw-Hill, 2010), 3rd ed.

[24] S. Franssila, *Introduction to Microfabrication* (Wiley, 2010), 2nd ed.

[25] G. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practices* (Academic Press, Inc., 1989).

[26] *2011 Minerals Yearbook* (United States Geological Survey, 2011).

[27] L. Parratt, "Surface studies of solids by total reflection of x-rays," Phys. Rev. **95**, 359–369 (1954).

[28] S. Maier, *Plasmonics : fundamentals and applications* (Springer Science+Business Media LLC, 2007), 1st ed.

[29] H. Raether, *Excitation of Plasmons and Interband Transitions by Electrons* (Springer-Verlag, 1980).

[30] M. Stockman, "Nanoscience: Dark-hot resonances," Nature **467**, 541–542 (2010).

[31] N. Christensen, "The Band Structure of Silver and Optical Interband Transitions," Physica Status Solidi (b) **54**, 551–563 (1972).

[32] M. Ordal, L. Long, R. Bell, S. Bell, R. Bell, J. Alexander, R., and C. Ward, "Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Ni, Pd, Pt, Ag, Ti, and W." Appl. Opt. **22**, 1099–1119 (1983).

[33] P. Worthing and W. Barnes, "Coupling effciency of surface plasmon polaritons to radiation using a corrugated surface; angular dependence," Journal of Modern Optics **49**, 1453–1462 (2002).

[34] C. Cojocariu and P. Rochon, "Light-induced motions in azobenzene-containing polymers," Pure and Applied Chemistry **76**, 1479–1797 (2004).

[35] C. Paik and H. Morawetz, "Photochemical and Thermal Isomerization of Azoaromatic Residues in the Side Chains and the Backbone of Polymers in Bulk," Macromolecules **5**, 171–177 (1972).

[36] T. Todorov, L. Nikolova, and N. Tomova, "Polarization holography. 1: A new high-efficiency organic material with reversible photoinduced birefringence," Applied Optics **23**, 4309–4312 (1984).

[37] P. Rochon, E. Batalla, and A. Natansohn, "Optically induced surface gratings on azoaromatic polymer films," Applied Physics Letters **66**, 136–138 (1995).

[38] D. Kim, S. Tripathy, L. Li, and J. Kumar, "Laser-induced holographic surface relief gratings on nonlinear optical polymer films," Applied Physics Letters **66**, 1166–1168 (1995).

[39] P. Rochon, J. Mao, A. Natansohn, and E. Batalla, "Optically induced high efficiency gratings in azo polymer films," Tech. rep., Queen's University Kingston (Ontario) Department of Chemistry (1994).

[40] X. Jiang, L. Li, J. Kumar, D. Kim, V. Shivshankar, and S. Tripathy, "Polarization dependent recordings of surface relief gratings on azobenzene containing polymer films," Applied Physics Letters **68**, 2618–2620 (1996).

[41] D. Kim, L. Li, X. Jiang, V. Shivshankar, J. Kumar, and S. Tripathy, "Polarized laser induced holographic surface relief gratings on polymer films," Macromolecules **28**, 8835–8839 (1995).

[42] D. Cumming, I. Khandaker, S. Thoms, and B. Casey, "Efficient diffractive optics made by single-step electron beam lithography in solid pmma," Journal of Vacuum Science & Technology B **15**, 2859–2863 (1997).

[43] T. Faivre (2014). Personal Communication.

[44] F. L. Pedrotti, L. S. Pedrotti, and L. M. Pedrotti, *Introduction to Optics* (Pearson Education, Inc., 2007).

[45] E. Hecht, *Optics* (Addison Wesley, 2002), 4th ed.

[46] V. Dhaka, T. Haggren, H. Jussila, H. Jiang, E. Kauppinen, T. Huhtio, M. Sopanen, and H. Lipsanen, "High quality gaas nanowires grown on glass substrates," Nano Letters **12**, 1912–1918 (2012).